ABSTRACT

Title of Dissertation: STRUCTURAL AND BIOLOGICAL STUDIES OF ATYPICAL RIO1 KINASE

Irene Njeri, Doctor of Philosophy, 2011

Dissertation Directed By: Assistant Professor Dr. Nicole LaRonde-LeBlanc
Department of Chemistry and Biochemistry

Rio1 kinase is an atypical protein kinase that is essential for the proper maturation of the 40S subsunit during ribosome biogenesis. Ribosome biogenesis is linked to cell cycle progression and studies have shown that the process is upregulated in both normal and abnormal highly proliferative cells in order to meet cell demands. In an effort to regulate cell proliferation especially in abnormal dividing cells, the work described here takes the approach of modulating ribosome biogenesis by targeting Rio1 kinase. Several small molecules were tested for binding to afRio1 and toyocamycin, an adenosine analog, had the highest binding affinity. A co-crystal structure of afRio1 with toyocamycin bound was determined at 2.0Å revealing occupancy of toyocamycin in the ATP binding pocket. Steady state kinetic studies revealed that inhibition occurred via both competitive and non-competitive mechanisms. This analysis also revealed that Rio1 accesses multiple oligomeric states and autophosphorylation reduces formation of oligomers. Autophosphorylation may therefore modulate Rio1 enzymatic activity through regulating oligomerization.

Additional inhibition studies revealed that afRio1 was susceptible to substrate inhibition at high concentrations of ATP.

The crystal structure of human Rio1 kinase (RioK1) was also determined, providing the first structure of a eukaryotic Rio1 kinase. This work was complemented with biochemical studies that revealed RioK1 shared similar trends as afRio1 including affinity for toyocamycin, formation of oligomers in the presence of different ligands and substrate inhibition at high concentrations of ATP suggesting that both afRio1 and RioK1 may share the same substrate inhibition mechanism.

In addition to work done on Rio1 kinase, structural determination of EspB, a *Mycobacterium tuberculosis* virulence protein involved in ESX-1 secretion system, is reported.

STRUCTURAL AND BIOLOGICAL STUDIES OF ATYPICAL RIO1 KINASE

By

Irene Njeri

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:

Professor Steve Rokita, Chair
Professor Nicole LaRonde-LeBlanc, Advisor
Professor Dorothy Beckett
Professor Kwaku Dayie
Professor Roy Mariuzza, Dean's Representative

Acknowledgements

First and foremost I would like to thank my advisor, Dr. Nicole Laronde-Leblanc. Her patience, support and encouragement have been fundamental through out my PhD career and most importantly for the opportunity that she gave me to continue on with my PhD career. She's nurtured me from an organic chemist mindset to a structural biologist. I'll forever be grateful for her mentorship, opportunity and continued support.

I would also like to thank my fellow lab members, both present and past especially Dr. Seth Thomas for his help, support and constructive criticism. I would also like to thank the Center of Biomolecular Structure Organization members for the consultations, conversations and laughter. They provided a friendly and relaxed atmosphere.

Thanks to my committee members, Dr. Dayie for his daily conversations and encouragement. His morning smile made some of my grumpy mornings a little brighter and hopeful. Dr. Beckett for your support, interest in my research and for assistance with the analytical centrifuge that was integral to my research. Dr. Roy Mariuzza for dedicating his time to serve as my Dean's representative.

Thanks to my family for your support, encouragement and love. My mom for her strength, encouragement, and love. Without her I would not be here and would have never embarked on this journey.

I would also like to thank Paulette Frazier. She's an amazing individual who was willing to put up with me for three years. Her door was always open for a conversation and her patience was immense.

Finally I'd like to thank, Lauren your love, patience, support and care has been much appreciated. When I lost track of my vision and future you always had a way of reminding me who I am, where I came from and where I want to go.

Table of Contents

List of Tables

List of Figures

ABBREVIATIONS

Å	Angstrom
ADE	Adenosine
ADP	Adenosine Diphosphate
AIDs	Acquired Immune deficiency syndrome
Ala	Alanine
APE	Activation loop
APKs	Atypical Protein Kinases1
APS	Advanced Photon Source
Arg	Arginine
Asp	Aspartate
Asn	Asparagine
ASU	Asymmetric Unit
ATP	Adenosine 5'-Triphosphate
BIS tris	Bis (2-hydroxyethyl)-amino-tris(hydroxymethyl)-methane
$CaCl_2$	Calcuim Chloride
CAPS	*N*-cyclohexyl-3-aminopropanesulfonic acid
$Ca(OAc)_2$	Calcium acetate
CCP4i	Collaborative Computational Project Number 4 Interface
CFP-10	Culture Filtrate Protein-10
CK2	Casein Kinase 2
COOT	Crystallographic Oject-Oriented Toolkit
DNA	Deoxyribonucleic acid
E.coli	Escherichia coli
Epks	Eukaryotic protein kinases
ESAT-6	Early secretory antigenic target-6
EspB	ESX-1 substrate protein B
ESX-1	ESAT-6 secretion system 1
FPLC	Fast protein liquid chromatography
G1	Gap 1
Gln	Glutamine
Glu	Glutamate
Gly	Glycine
GTP	Guanosine 5'-Triphosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His	Histidine
HIV	Human Immunodeficiency Virus
HP	High Performance
HYSS	Hybrid Substructure Search
Ile	Isoleucine
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K	Kelvin
K_2HPO_4	Potassium phosphate dibasic
K_d	Dissociation constant
kDa	Kilodaltons

K_m	Affinity of enzyme for substrate
L	Liter
LB	Lysogeny Broth
Leu	Leucine
Li_2SO_4	Lithium sulfate
Lys	Lysine
M	Molar
MBP	Myelin Basic Protein
MDR-TB	Multi-drug resistant Tuberculosis
Met	Methionine
mg	Milligrams
Mg	Magnesium
$MgCl_2.2H_2O$	Magnesium Chloride dihydrate
$Mg(COO)_2$	Magnesium acetate
mL	Milliliter
mM	Millimolar
M.Marinum	Mycobacterium Marinum
Mn	Manganese
MOLREP	Molecular Replacement
Mtb	Mycobacterium tuberculosis
Mw	Molecular Weight
NaCl	Sodium Chloride
NaH_2PO_4	Sodium dihydrgogen phosphate
NE-CAT	Northeastern Collaborative Access Team
nm	nanometer
nM	Nanomolar
OD	Optical Density
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase Chain Reaction
PDB	Protein Database Bank
PEG	Polyethylene glycol
Phe	Phenylalanine
PHENIX	Python-based Hierarchical Environment Integrated Xtallograhy
PKA	Protein Kinase A
PKI	Protein Kinase Inhibitor
PISA	Protein Interfaces, Surfaces and Assemblies
pmol	Picomolar
PRMT5	Protein Arginine Methyltransferase 5
Pro	Proline
RD1	Region of Difference 1
Rio	Right Open reading frame
RMSD	Root Mean Square Deviation
RNA	Ribonucleic Acid
RNP	Ribonucleoprotein
rpm	Revolution per minute
rRNA	Ribosomal ribonucleic acid

SAD	Single Anomalous Dispersion
SDS	Sodium Dodecyl Sulfate
Sec	Second
SEC	Size exclusion chromatography
Se-Met	Selenomethionine
Ser	Serine
SER	Surface Entropy Mutagenesis
S_N2	Bimolecular nucleophilic substitution
snoRNP	Small nucleolar ribonucleic acid
TB	Tuberculosis
TEV	Tobacco Etch Virus
Thr	Threonine
TOYO	Toyocamycin
TRIS	Tris(hydroxymethyl)aminomethane
tRNA	Transfer ribonucleic acid
Tyr	Tyrosine
µg	Microgram
µL	Microliter
µM	Micromolar
Val	Valine
V_{max}	Maximum catalytic rate
V_o	Initial Velocity
XDR-TB	Extensively drug-resistant Tuberculosis

Chapter 1: Introduction

Protein kinases are involved in various cellular processes such as regulation of gene expression, DNA transcription, signal transduction, cell cycle progression and ribosome biogenesis (*1-5*). Kinases' involvement in numerous cellular processes implicate them in diseases such neurodegenerative disorders, cystic fibrosis, HIV/AIDs and cancer (*6-9*). These enzymes catalyze the phosphorylation reaction that entails the transfer of a γ- phosphate of purine nucleotide triphosphates (ATP and GTP) to protein residues such as serine, threonine and tyrosine (*10*). Two phosphorylation mechanisms have been proposed; the S_N2 reaction and general base catalysis (*11, 12*). S_N2 reaction involves a concerted attack of a nucleophile (the hydroxyl group on the residues) onto the γ-phosphate of purine nucleotide triphosphates (Figure 1.1). This mechanism was first supported by isotopic and crystallographic studies of a ternary complex of Protein Kinase A (PKA) complexed with ATP/Mg^{2+} and its inhibitor, PKI (*13-15*). On the other hand the general base catalysis is supported by crystallographic studies of PKA complexed with ATP and Kemptide, a PKA substrate (*12*). It involves a conserved aspartate residue that functions as a catalytic base by deprotonating the hydroxyl group to promote nucleophilic attack on the γ- phosphate (*12*). The S_N2 reaction is the more acceptable mechanism given that in general base catalysis deprotonation of the hydroxyl group results in a negatively charged oxygen that would cause charge repulsion with the negatively charged phosphates of ATP.

Kinases are among the largest protein superfamilies with over 518 members identified in the human genome (*10, 16, 17*). These kinase members are classified as

either eukaryotic protein kinases (ePKs) or atypical protein kinases (aPKs) and they both contain a kinase signature that's comprised of conserved residues required for kinase activity. However, aPKs are not significantly related to ePKs in sequence and there are more ePKs members compared to aPKs. Currently they are 40 atypical protein kinase members in thirteen different families including the RIO family (*16, 18*).

1.1 Atypical RIO Kinase Family

Evolutionary studies indicate a shared ancestry between ePKs and RIO kinases with an indication that RIO kinases predated ePKs (*19, 20*). The RIO family was first discovered through genomic studies in budding yeast in 1996 that identified the **ri**ght **o**pen reading frame 1 (RIO1) (*21*). The following year, the RIO family was characterized based on studies of Rio1 that showed that it was a serine kinase that was expressed constitutively at low levels in yeast cells and had a distant sequence similarity with ePKs (*22, 23*). Since then progress has been made in this family with the identification of four subfamilies, bacterial Rio, Rio1, Rio2 and Rio3. In addition crystal structures of Rio1 and Rio2 from *Archaeoglobus fulgidus* that were integral in defining the RIO catalytic domain, have been determined. *(24-26)*. Rio1 and Rio2 are present from archaea to humans whereas Rio3 is only found in multi-cellular eukaryotes (*27, 28*). These subfamilies are distinguishable by their subfamily-specific N-terminal domains that are found outside the RIO domain (Figure 1.3) (*24, 27*). Rio1 has two additional alpha helices (α1 and α2) on the N-terminal whereas Rio2 has a winged-helix domain, a domain which is mostly found to participate in DNA

binding (Figure 1.2) (*24, 26*). On the other hand, the N-terminal structural feature of Rio3 has not yet been identified due to lack of a crystal structure. Primary sequence analysis of the RIO catalytic domain shows that Rio1 and Rio3 are most closely related in the RIO domain, consistent with the idea that Rio3 arose in metazoa from gene duplication of Rio1 (Figure 1.3) (*27*).

1.2 RIO catalytic Domain.

RIO catalytic domain has a 3-dimensional fold consisting of an N-lobe composed of a β-sheet adjacent to a single α helix (αC) and a helical C-lobe (Figure 1.4). The RIO catalytic domain has subdomains with conserved residues required for binding and orientation of ATP as well as transfer of the γ-phosphate to the substrate (*24-26*). These subdomains include a nucleotide binding loop (p-loop) that binds the phosphates on ATP; a hinge region found between the N-lobe and the C-lobe that binds the adenine moiety via hydrophobic and hydrogen bonding interactions; a catalytic loop that contains a catalytic aspartate and a conserved asparagine required for phosphoryl transfer; and a metal-binding loop (DFG loop) that has a conserved aspartate residue for positioning divalent cations (*24, 29*) (Figure 1.4 A-B). The RIO domain lacks an activation loop (APE loop) that regulates kinase activity and substrate binding subdomains (subdomain X and XI) that are present in the ePK catalytic domain. This suggests that RIO proteins bind and interact with substrates differently from ePKs. Another observable difference between the ePK and RIO catalytic domains is the difference in conformation of bound ATP, specifically the distance between the γ-phosphate and the catalytic aspartate. For instance in Protein

Kinase A, an example of ePKs, this distance is 3.8 Å compared to 5.1 Å in *Archaeoglobus fulgidus* Rio1 suggesting a possible difference in catalytic mechanism (*24, 30*).

1.3 Rio1 Kinase: Background and Roles

Rio1 kinase is the founding member of the RIO family and based on autophosphorylation and phosphoamino acid analysis, this protein is classified as a serine kinase (*24, 31*). Rio1 is an essential gene in the yeast species *Saccharomyces cerevisiae* and it is required for cell cycle progression, chromosome maintenance and ribosome biogenesis (*18, 32, 33*). During cell cycle progression Rio1 kinase is needed for exit from both the G1 phase and mitosis (*32*). Cells depleted of Rio1 are enlarged and found in clusters indicative of arrest in the G1 phase and some cells have a single DNA mass at the bud neck as well as short spindles indicative of arrest at the metaphase during mitosis (*32*).

Rio1's involvement in chromosome maintenance occurs during mitosis, where Rio1 is needed for chromosome stability and nuclear segregation. This is evident in cells depleted of Rio1 that have faulty DNA replication, incomplete DNA repair and defects in sister chromatid segregation (*32*). In addition to cell cycle progression and chromosome maintaince, studies in *Saccharomyces cerevisiae,* have identified Rio1 kinase as a non-ribosomal protein involved in the maturation of 18S rRNA during ribosome biogenesis (*34*).

1.4 Ribosome Biogenesis and Rio1's involvement.

A lot of progress has been made in the study of ribosomes from 1958 when George Palade used electron microscopy to identify the varying patterns of the ribosome, for which he won a nobel prize, to the current crystal structure of the 80S eukaryotic ribosome solved in 2010 (*35-43*). Ribosomes are burdened with correctly and efficiently producing proteins in the cell making ribosome biogenesis a highly complex and well-coordinated process in both prokaryotes and eukaryotes. In eukaryotes, ribosome biogenesis has been well studied in yeast cells where the process starts in the nucleolus with the synthesis of the precursors 35S and 5S pre-rRNAs by RNA polymerase I and III respectively (*44, 45*). The 35S precursor forms a large ribonucleoprotein (RNP) complex with both ribosomal and non-ribosomal proteins. This complex is processed to the 60S subunit (5.8S, 25S rRNA) and the 40S subunit (18S rRNA) (*46*). The rRNA precursors are extensively modified and their maturation as well as assembly into ribosomal subunits involves more than 170 proteins (*45*). These proteins include endo- and exonucleases, ribosomal and non-ribosomal factors, chaperones, RNA helicases and small nucleolar ribonucleoprotein particles (snoRNPs) (*44, 45*). Some of the modifications include pseudouridylation, 2'O-methylation and cleavage of pre-rRNAs. A brief description of the process begins with pseudouridylation and 2'O-methylation of the 35S pre-rRNA. The 35S pre-rRNA is then cleaved at A_0 and A_1 sites resulting in the 32S pre-rRNA that is further cleaved at the A_2 site producing $27SA_2$ and 20S pre-rRNAs (Figure 1.5) (*44, 45*). The $27SA_2$ undergoes further cleavages to produce mature 25S and 5.8S

rRNAs that are assembled in 60S subunit. The 20S pre-rRNA undergoes cleavage at the D-site to produce mature 18S rRNA that is assembled in the 40S subunit (Figure 1.5). Rio1 kinase is involved in this cleavage at the D-site and cells depleted of Rio1 have an accumulation of the 20S pre-rRNA (*33, 34*).

A recent report has shown the involvement of human Rio1 (RioK1) in ribosome biogenesis through complex formation with human protein arginine methyltransferase 5 (PRMT5) (*47*). In this complex, RioK1 recruits nucleolin, a protein involved in multiple functions in ribosome biogenesis such as binding to nascent rRNA and providing cytoplasmic stability (*48*). Once nucleolin is recruited to the PRMT5 complex, it undergoes dimethylation that is necessary for modulating nucleolin's interactions with rRNA (*47*).

1.5 Linking Ribosome Biogenesis to Cell Proliferation.

Ribosome biogenesis is upregulated in both normal and abnormal proliferative cells. Upregulation of ribosome biogenesis is necessary for increased protein production and production of new ribosomes for new daughter cells. A characteristic feature of highly proliferating cells such as cancer cells is an enlarged nucleolus. The nucleolus is the compartment where rRNA processing occurs (*49-51*). This enlargement is a result of increased levels of rRNA transcription, rRNA processing and increased expression of ribosomal proteins (*51*). Therefore, there is a need to regulate cell proliferation, especially in abnormal dividing cells by possibly modulating the rate of ribosome biogenesis. This approach takes on the idea of targeting and inhibiting either ribosomal or non-ribosomal proteins. It has previously

been shown that blocking ribosome biogenesis inhibits cell proliferation but not cell growth. This is well illustrated in mice, where deletion of the ribosomal protein S6 in the liver results in reduced cell proliferation but the liver is able to grow in response to nutrient stimuli (*52*). Bolstered by this data and further evidence that RioK1 is overexpressed in colon cancer cells, some of the work described in this thesis seeks to address the possibility of inhibiting Rio1 by carrying out both biochemical and crystallographic studies (*53*).

1.6 Protein x-ray crystallography: The phase problem and the solutions.

Protein crystallography dates back to 1934 when J.D. Benal and Dorothy Crowfoot measured diffraction from hydrated pepsin crystals. The first crystal structure solved was the sperm whale myoglobin in the late 1950s by Max Perutz and Sir John Cowdey Kendrew for which they both received the Nobel Prize in chemistry in 1962 (*54*). Since then over 59,007 protein structures have been solved and deposited to the protein data bank (PDB). Along with the increased number of solved protein structures, x-ray crystallography methods have also advanced with better detectors and improved software for structure determination. However, one obstacle that crystallographers face is the phase problem that arises during data collection when the crystal diffracts x-ray beams that are recorded as reflections. These reflections are treated as complicated waves and the ultimate goal of a crystallographer is to obtain an electron density map that reveals the molecular shape of the protein. In order to obtain the electron density map, a periodic function that is

defined by Equation 1 is used to transform recorded reflections to an electron density map.

$$\rho(x,y,z) = \frac{1}{V}\sum\sum\sum | F_{hkl} | e^{-2\pi i(hx+ky+lz-\alpha)}$$

Equation 1: Electron density equation ρ(x,y,z). A sum of the reflections' amplitude ($|F_{hkl}|$), frequencies (h,k,l) and phases (α).

To solve this periodic function, the amplitudes, frequencies and phases have to be determined. Only two of these parameters are directly acquired during data collection, the amplitude that is represented by $|F_{hkl}|$ and the frequencies represented by h, k, l, in the equation. Unfortunately the phases cannot be determined directly and methods for acquiring phases have been developed that include molecular replacement and single wavelength anomalous dispersion (SAD). Molecular replacement uses phases from a previously solved crystal structure as initial phases for the new protein. For molecular replacement to be successful, the sequence identity between the new protein and the known protein, also known as the search model, has to be at least 30%. Unlike molecular replacement, SAD phasing relies on heavy atoms to obtain phases for the new protein. The heavy atoms can be introduced by either soaking in ionic solutions or as modified amino acids during expression. The most common heavy atom is selenium and it is incorporated into the protein during expression in media that has methionine replaced with selenomethionine (Se-Met). Phasing by this method takes advantage of the heavy atom's ability to violate or

break Friedel's law that requires intensities of symmetry related reflections to be equal i.e. $I_{hkl}=I_{(-h,-k,-l)}$. Once data is collected, the heavy atoms are located by Patterson maps and the initial phases are calculated and then improved by density modification (*55-57*). Both phasing methods have been used in the work that will be described in chapter 2, 4 and chapter 6.

1.7 Objectives

The ultimate goal is to regulate cell proliferation by possibly modulating the rate of ribosome biogenesis. Towards this ultimate goal, the focus of my research was to target the non-ribosomal protein, Rio1 by asking the following questions:

1. What and where do we begin the search for potential Rio1 inhibitors that will aid in designing more specific inhibitors?
2. If Rio1 potential inhibitors do exist and the mode of inhibition?
3. Since initial studies will be carried out with *Archaeoglobus fulgidus* Rio1 (afRio1), are the properties determined conserved in and applicable to eukaryotic Rio1 kinase?
4. And finally, what is the structure of eukaryotic Rio1 and how does it compare to the archaeal structure?

To address these questions, I solved the crystal structure of *Archaeoglobus fulgidus* Rio1 (afRio1) with a potential inhibitor bound. I will describe this complex in chapter 2. In chapter 3, I will describe as well as characterize the mode of

inhibition by examining the steady state kinetics and in chapter 4, I will describe the work that went towards determining the structure of the human Rio1 (RioK1) and relevant biological studies. Chapter 5 will focus on functional and mutational studies of afRio1.

1.8 Additional work: EspB

In addition to my main project, I took on another project that entailed structural determination of EspB, a protein that is involved in the virulence of Mycobacterium tuberculosis (*MtB*). This structure was integral to my protein crystallography experience. Expression, purification and structural determination will be described in detail in chapter 6.

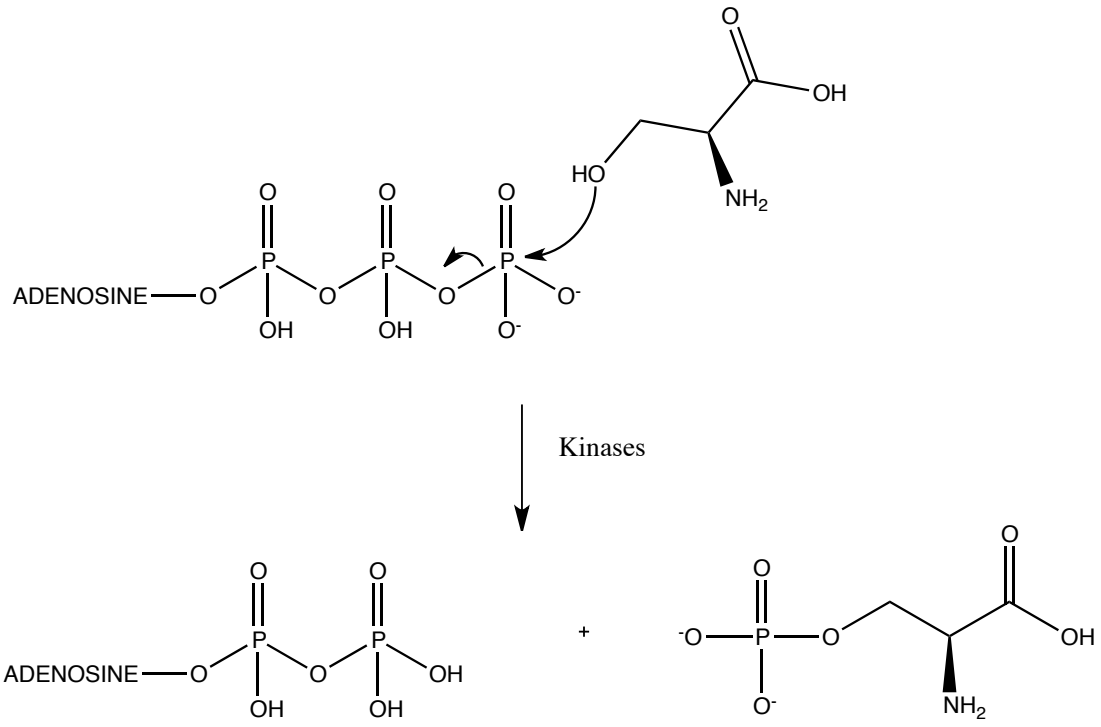


Figure 1.1: **Phosphorylation mechanism**. S_N2 mechanism of the phosphorylation reaction.

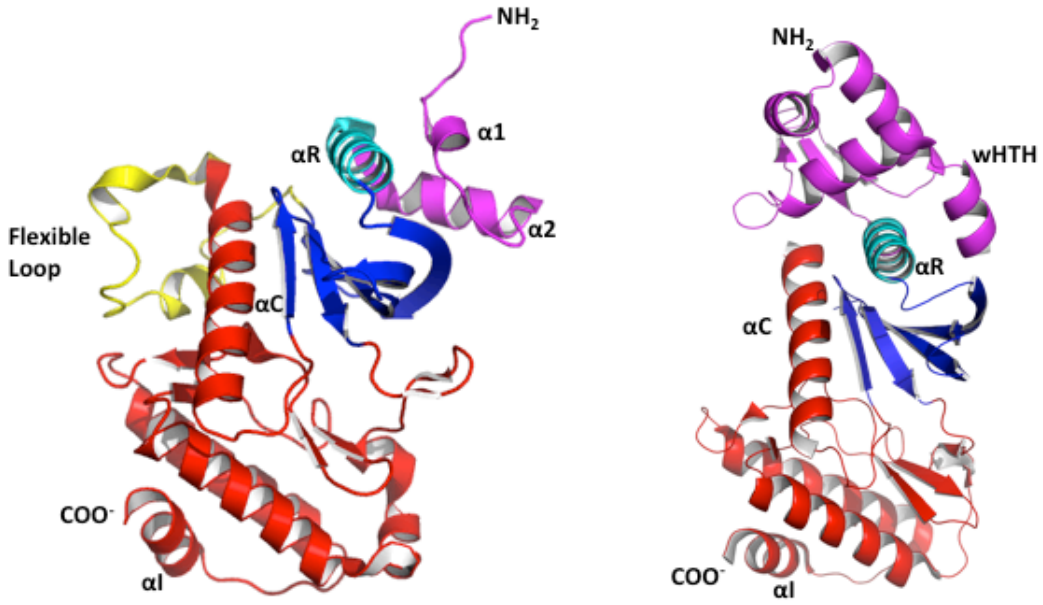


Figure 1.2: Crystal Structures of Rio1 and Rio2 from *Archaeoglobus fulgidus*. **A.**Rio1with the additional alpha helices (α1 and α2) in magenta. **B.** Rio2 with the winged helix domain (wHTH) in magenta. In both structures the Rio domain starts at αR and ends at αI.

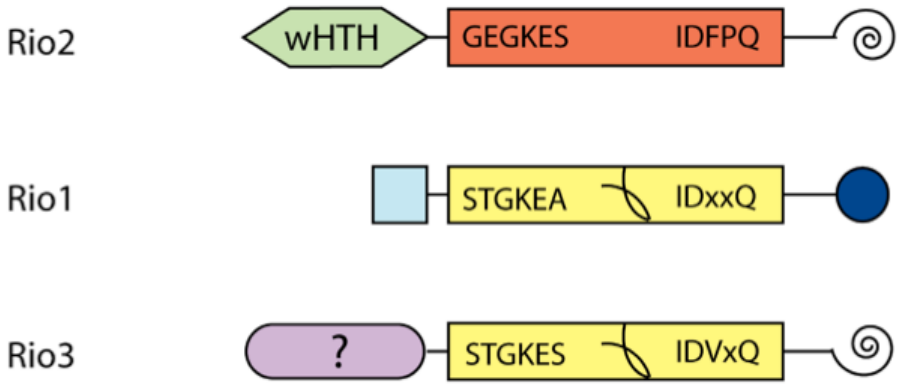


Figure 1.3: Schematic of primary structure of Rio kinases. Adapted from LaRonde-LeBlanc (*27*). Primary sequence of Rio1, Rio2 and Rio3 showing the different N terminal domain and the sequence variations in the RIO domain that's in yellow in Rio1 and Rio3 and orange in Rio2. Rio1 and Rio3 have similar conserved residues in the RIO domain.

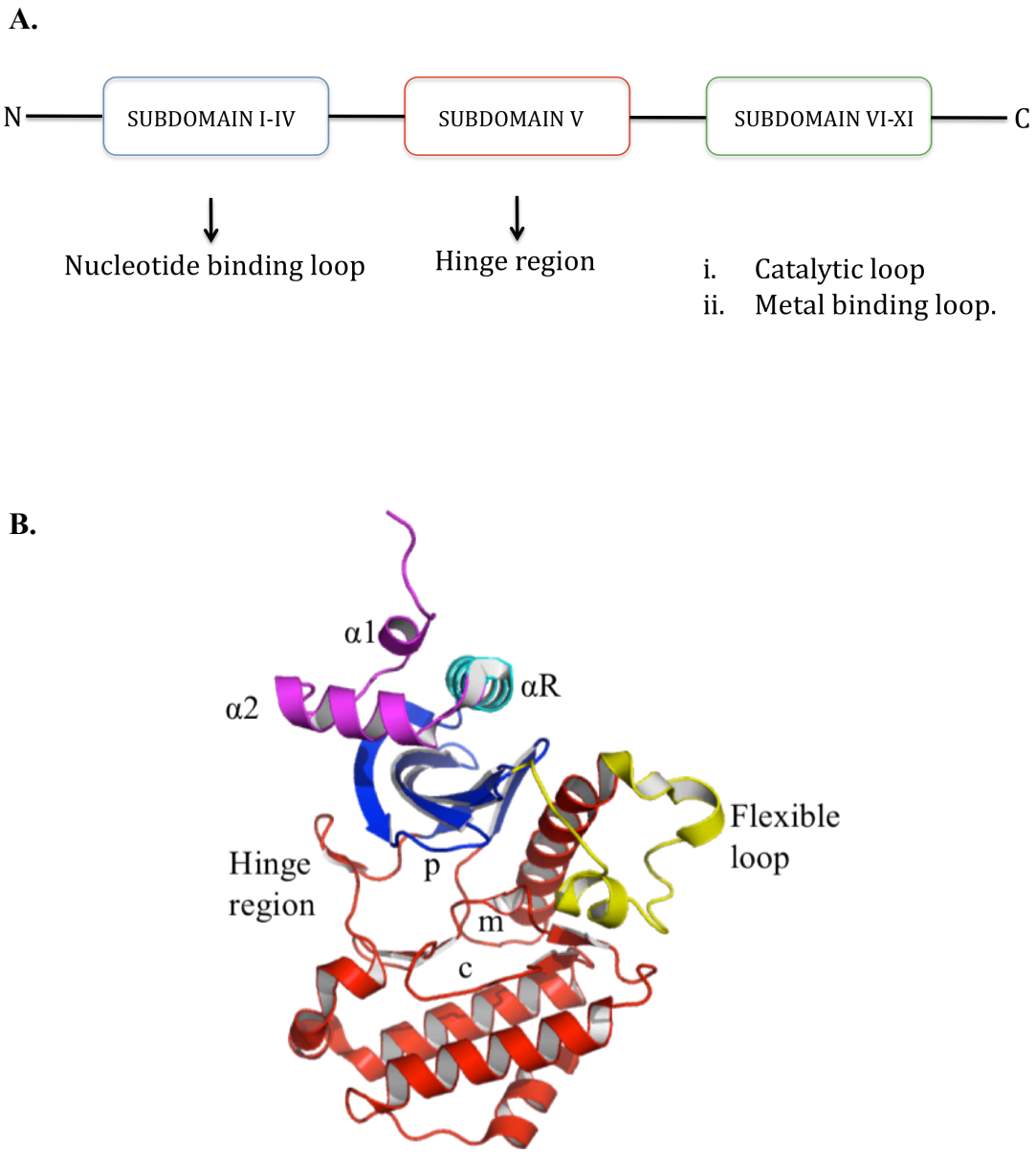


Figure 1.4: RIO catalytic domain. **A.** Simple schematic of the Rio domain. **B**. Archaeal Rio1 (afRio1) structure (PDB ID: IZTF) with the RIO domain labeled; p is the nucleotide-binding loop, m is the metal binding loop and c is the catalytic loop.

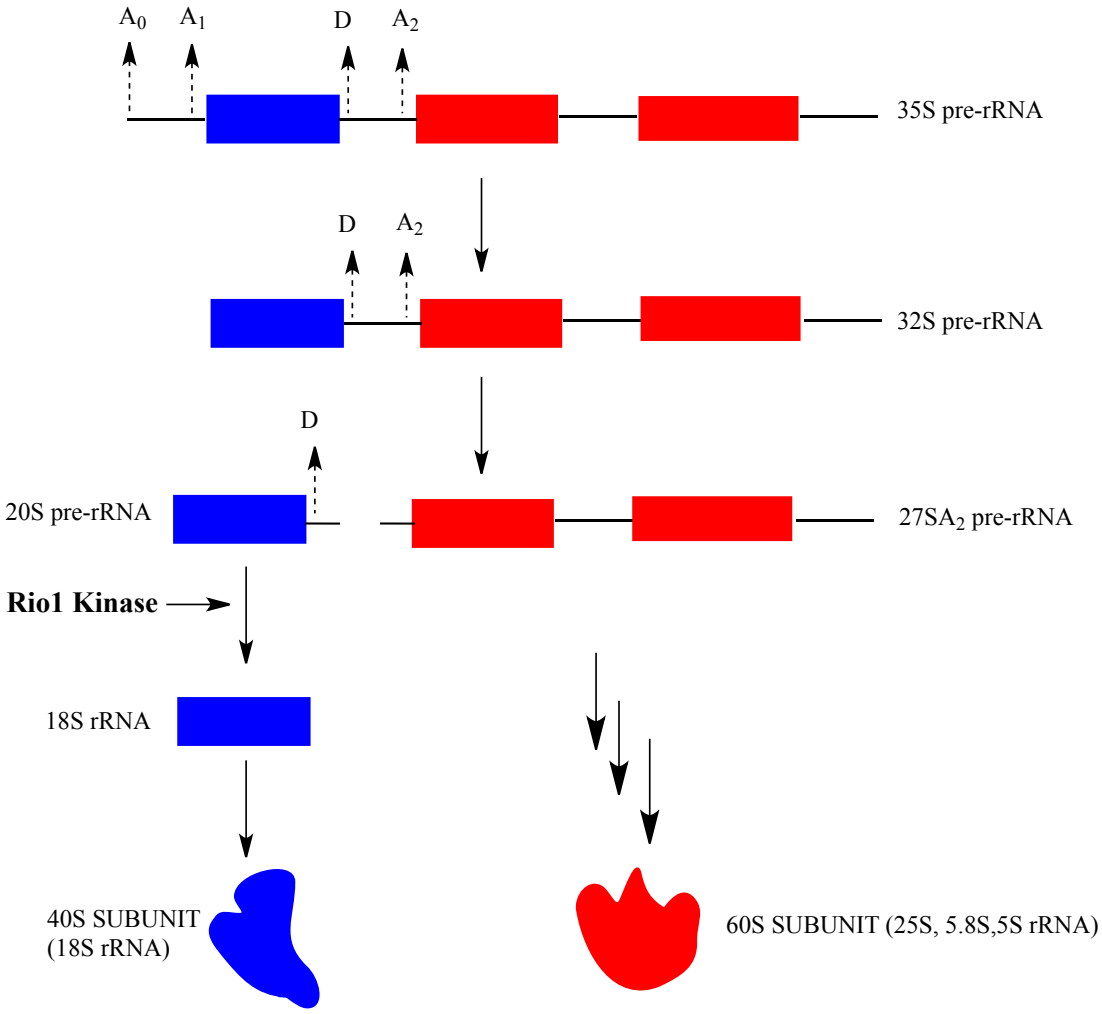


Figure 1.5: Ribosome Biogenesis. Simple schematic of ribosome biogenesis starting with large transcript 35S pre-rRNA that undergoes several modifications and cleavages to produce the mature rRNA that are incorporated in the respective subunits.

Chapter 2: Structure Determination of Toyocamycin Bound to *Archaeoglobus fulgidus* Rio1 Kinase (afRio1).

2.1 Introduction.

In an effort to provide a means of modulating the rate of ribosome biogenesis with the ultimate goal of regulating cell proliferation, we decided to target Rio1, as it is involved in the maturation of 18S rRNA (*34*). The first step was to screen and identify small molecules that bind to afRio1 with a higher affinity than the native ligand, ATP/Mg^{2+}, followed by co-crystallization studies in order to characterize the interactions.

A previously solved crystal structure of afRio1 with ATP bound revealed an accessible pocket in the ATP binding site where N7 of the adenine moiety is positioned (Figure 2.2) (*24*). Based on this structure three adenosine analogs, toyocamycin and sangivamycin and 7-methylinosine were screened for binding to afRio1. Toyocamycin has a carbon and a cyano group at the N7 position whereas sangivamycin has a carbon and an amide group at the same position (Figure 2.3). 7-methylinosine has a methyl group at the N7 position and a hydrogen bond acceptor (a carbonyl group) at the C6 position that is not optimal for binding to afRio1, as is evident in previously solved crystal structures of afRio1 with bound ATP, ADP and adenosine bound (Figure 2.3) (*24*). In these structures the N6 position has a hydrogen donor (an amine group) that hydrogen bonds with the peptidyl carbonyl of Glu 148 (*24*). In order to determine the affinity of these small molecules for afRio1, a thermofluor assay was carried out. In this assay, the protein is slowly heated up and

as the protein unfolds the inner hydrophobic core residues are exposed, allowing for hydrophobic fluorescent dyes to bind (Figure 2.4) (*58*). A change in fluorescent intensity is monitored as temperature increases and the melting temperature (T_m) is determined as the inflection point of the melting curve. Thermofluor assays monitor the shift in T_m by comparing the T_m of unbound protein, the control, to the T_m of protein bound to a ligand. A positive shift in T_m is indicative of strong binding and stabilization whereas a negative shift in T_m is indicative of destabilization of the protein by the ligand.

Toyocamycin was identified as a tight-binding small molecule. This chapter will describe the journey from expression and purification of afRio1 to structural determination of toyocamycin bound to afRio1 as an initial step to designing Rio1-specific inhibitors.

2.2 Toyocamycin background

Toyocamycin, also known as 7-cyano-7-deazaadenosine, was first isolated in 1956 from *Streptomyces Toyocaensis*, a gram-positive bacteria (*59, 60*). Toyocamycin was discovered as a pyrrolopyrimidine antiobiotic and later identified as an antitumor agent that can also serve as a precursor to other antitumor agents such as sangivamycin (Figure 2.6) (*61*). Studies as early as 1968 showed toyocamycin suppressed the formation of 28S and 18S rRNA, the mature mammalian rRNAs in the large and small subunit respectively (*62*). These studies were carried out in several cancer cell lines including HeLa cells, colon carcinoma cells and chick embryo cells infected with MC29 avian Leukosis virus (*63-65*). Other studies investigating the

effect of low concentrations of toyocamycin on normal mammalian cells showed interference with the maturations of precursors ranging from 45S to 32S pre-RNAs (*66*). These precursors accumulated in the nucleoli, the compartment where rRNA processing begins, resulting in the lack of formation of 28S and 18S rRNA. In addition to inhibiting rRNA processing, toyocamycin has been shown to inhibit several kinases including phosphatidylinositol kinase, rhodopsin kinase and adenosine kinase from Mycobacterium tuberculosis (*67-70*). In order to test the binding affinity of toyocamycin to afRio1 and characterize the binding interaction the following experimental procedures were carried out.

2.3 Experimental Procedures:

2.3a Expression of afRio1.

AfRio1 expression vector (cloned by my advisor, Dr. LeBlanc) with a T7 promoter, N-terminal 6X histidine tag, Tobacco Etch Virus (TEV) was transformed into Rosetta DE3 *E.coli* cells (Novagen). These cells have tRNAs for rare codons (AUA, AGG, AGA, CUA, CCC, CGG, and GGA) and a pLysS plasmid that produces lysozyme to suppress basal expression from T7 promoter. The cells were plated on Lysogeny Broth (LB) Agar plates containing 100 μg/mL ampicillin and 34 μg/mL chloramphenicol and incubated at 37°C for 16 hours. Choice of antibiotics was based on ampicillin resistance provided by the expression vector and maintenance of pLysS plamid that requires chloramphenicol. A single colony was inoculated in 40 mL of LB media containing 34 μg/mL chloramphenicol and 100 μg/mL carbenicilin, an analog of ampicillin that is more resistance to degradation. This LB media was

incubated overnight at 37°C while shaking at 250 rpm. The following day, 40 mL of overnight culture was added to 4L of LB media containing 100 μg/mL ampicillin and 34 μg/mL chloramphenicol and incubated at 37°C while shaking at 250 rpm. When the OD_{600} reached 0.6-0.8, within the mid-log phase, these cells were induced with 1 mM Isopropyl β-D-1-thiogalactopyranoside (IPTG) and expression proceeded at 20°C for 18 hours. The cells were centrifuged at 6,000 rpm and 4°C for 30 minutes using a Sorvall centrifuge (Dupont) with a SLC 4000 rotor. The bacterial pellet was stored at -80°C.

2.3b Purification of afRio1

The frozen bacterial cell pellet was resuspended in lysis buffer containing 50 mM Tris pH 8.0, 200 mM NaCl, 0.2% β-mercaptoethanol, 0.1 mg/mL DNase1 (Roche) and 0.3X Bugbuster (Novagen) and gently stirred on ice for 30 minutes. The lysate was transferred to pre-cooled centrifuge tubes that were placed in a 75°C waterbath for 15 minutes to denature *E.coli* proteins and immediately centrifuged at 4°C and 15,000 rpm for 20 minutes using a Beckman Coulter ultracentifuge with a 45 Ti rotor. The supernatant was filtered and loaded onto a pre-equilibrated 5 mL HisTrap™ HP column (GE healthcare) attached to a Fast Protein Liquid Chromatography (FPLC) system at a flowrate of 2.5 mL/min. The column was washed with 10 column volumes of lysis buffer to remove any unbound protein. Bound afRio1 protein was subjected to an imidazole gradient of 100 mM to 1 M and afRio1 eluted between 200 mM to 300 mM imidazole. Partially purified fractions were pooled together and transferred to a dialysis membrane to which 1 mg/mL of

Tobacco Etch Virus (TEV) protease was added to cleave the histidine tag. AfRio1 was dialyzed overnight at 4°C into 50 mM Tris pH 8.0, 200 mM NaCl, 0.2% β-mercaptoethanol. After cleavage the protein was loaded onto a pre-equilibrated 5 mL HisTrap™ HP column (GE healthcare) and washed with lysis buffer. Cleaved protein eluted with the lysis buffer and was further purified by size exclusion chromatography using a Superdex™ 200 column (GE healthcare) that was pre-equilibrated in 10 mM Tris pH 8.0, 10 mM NaCl and 0.2% β-mercaptoethanol. All purification steps were monitored by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) for purity (Figure 2.1).

2.3c Thermofluor Assay

Purified afRio1 (1.0 mg/mL) was incubated at 25°C for 30 minutes with 1 mM toyocamycin, sangivamycin, 7-methylinosine and ATP/Mg^{2+} in a buffer containing 50 mM Tris pH 8.0, 75 mM NaCl and 2 mM $MgCl_2$. The control reaction was afRio1 with no ligand added. After incubation, Sypro-Orange™ fluorescent dye (Sigma-Aldrich) was added to the reactions and heated from 30°C to 98°C at a rate of 0.2°C per second using Bio-Rad Mini-Opticon™ thermal cycler. Measurements were done in triplicates and the T_m was determined as the inflection point of the melting curve. The change in T_m was calculated as the difference between the T_m of bound afRio1 and T_m of unbound afRio1. The T_m of unbound afRio1 was 76°C ±0.4°C whereas in the presence of toyocamycin a significant shift was observed with a T_m of 88°C ±0.1°C, corresponding to a 12°C shift. In the presence of ATP/Mg^{2+} or sangivamycin the T_m was 82°C ±1.2°C and 82°C ±1.0°C respectively, a 6°C shift for

both molecules (Figure 2.5). In the case of 7-methylinosine the T_m was 79°C ± 0.2 °C having the least shift in T_m of only 3°C (Figure 2.5). These results revealed that toyocamycin had the most significant shift and therefore was the most stabilizing small molecule bound to afRio1. In order to characterize the interaction between afRio1 and toyocamycin, a co-crystallization experiment was carried out.

2.3d Co-crystallization of afRio1 with toyocamycin.

Purified afRio1 (21 mg/mL) was incubated with 2 mM toyocamycin at room temperature for one hour allow for maximum binding in a buffer containing 10 mM Tris pH 8.0, 10 mM NaCl and 0.2% β-mercaptoethanol. This was followed by initial screening of crystals with commercially available screen matrices; Index (Hampton), Cyros (Qiagen), PEGs suite (Qiagen), Wizard I, II and III (Emerald Biosystems) in 96-well plates. The initial screens were setup using the Phoenix Liquid Handing system (Art Robbins) in three different protein to well solution ratios (1:2, 1:1, 2:1) using the sitting drop method. In this method, a drop of protein and well solutions (precipitants, salts and/or buffers) sit in a depression and equilibration occurs by vapor diffusion. The 96-well plates were incubated at 20°C and afrio1-toyocamycin crystals were observed after seven days in PEGs Suite screen (Figure 2.7). These crystals grew in a 1:2 protein to well solution ratio in 0.2 M $Mg(COO)_2$, 20% PEG 3350, pH of 8.0. The crystals were optimized by varying the both the salt concentration, $Mg(COO)_2$, from 0.15 M to 0.3 M and the precipitant, PEG 3350, from 14% to 32%. Optimization was done using the hanging drop method in 24-well plates and incubated at 20°C. This method entails having a drop of protein and well solution

mixture placed on a glass slide or seal that covers a well containing a solution of buffers, salts and precipitants. The drop "hangs" over the well and equilibration occurs via vapor diffusion.

Optimized crystals were observed after a week. Several of these crystals were flash frozen in their respective well solution and 20% (v/v) glycerol as a cryo-protectant to reduce radiation damage and prevent ice formation during data collection. The crystals were shipped to Argonne National Lab where data collection and processing was performed at NE-CAT beamline.

2.3e AfRio1-toyocamycin complex: Data collection, processing and refinement.

Data was collected at 100K at the NE-CAT beamline at Advanced Photon Source (APS), Argonne, IL, USA. The best diffraction was to 2.0 Å resolution and data was indexed, integrated and scaled using HKL2000 package program (*71*). The best diffracting crystal belonged to $P2_1$ space group and the unit cell dimensions were a= 52.81Å, b= 72.61Å, c= 60.56Å and $\alpha= 90.00^0$, $\beta= 90.17^0$, $\gamma= 90.00^0$. The overall redundancy was 4.0 with 97.7% data completion and an overall R_{symm} of 6.9% (Table 2.1). These three parameters were used to determine the quality of data collected. The redundancy is the average number of independent measurements of each reflection in a data set and an average value above 3.0 is generally acceptable. The R_{symm}, which is also referred to, as R_{linear} is the measure of agreement among independent measurements of symmetry-related reflections in a data set and a value below 18% is acceptable.

Molecular replacement was used to solve afRio1-toyocamycin structure and a

previously solved crystal structure of afRio1 (PDB ID: IZTF) was used as a search model to phase afRio1-toyocamycin. Molecular replacement was carried out using MOLREP as part of CCP4i and the model was subjected to multiple rounds of model building in COOT and several rounds of refinement using PHENIX and REFMAC5 that is part of CCP4i (*72-77*). Waters were added to the structure using PHENIX and the final model had a R_{work}/R_{free} of 21.6/25.5 (Table 2.2). The completely refined crystal structure contained two molecules of afRio1-toyocamycin complex per asymmetric unit (ASU) (Figure 2.8). This structure has been deposited to the Protein Data Bank (PDB ID: 3REA).

2.4 Results

2.4a Overall structure of afRio1-toyocamycin complex.

The overall structure of afRio1-toyocamycin complex consists of an N-lobe that has a α-helical region (α1, α2 and αR) as well as a twisted β sheet that contains the nucleotide binding loop (p-loop) and an α-helical C-lobe that has the catalytic and metal binding loops (Figure 2.8). This overall structure is consistent with the general features of kinases that contain an N-terminal comprised of a twisted β-sheet and an α-helical C-terminal (*10, 24*). The N-lobe and C-lobe are connected by a hinge region that buries the 7-cyano-7-deazaadenine ring of toyocamycin. This region buries the adenine ring of ATP in the afRio1-ATP complex (*24*).

AfRio1 kinase has a flexible loop consisting of residues 83 to 111 that house the autophosphorylation site, serine 108 (*24*). No electron density was observed for part of this flexible loop in the afRio1-toyocamycin complex; Chain A and B lack

density for residues 86-89 and 86-92 respectively. This is similar to the structure of ATP/Mg^{2+} bound to afRio1 for which no electron density is observed for residues 85-91 (*24*). The RMSD between the two molecules of afRio1-toyocamycin in the asymmetric unit is 0.024 $Å^2$ for 240 residues. This structure provides the first structural information linking toyocamycin to a ribosomal processing factor.

2.4b Interactions of toyocamycin bound to afRio1.

To generate a difference electron density map (Fo-Fc) to confirm the presence of toyocamycin, the afRio1-toyocamycin complex was refined in the absence of any ligand and the resulting structure factors and phases were used to generate the map. This revealed electron density for toyocamycin in the ATP binding site (Figure 2.9A). Toyocamycin is buried in a pocket that features hydrophobic contact with Val 63, Ile 55 and Ala 78 from the N-lobe, Pro 156, Met 147, Phe 149 and Ile 150 from the hinge, and Met 203 and Ile 211 from the C-lobe (Figure 2.9A). In addition, toyocamycin participates in ten hydrogen bond interactions. There is one direct interaction from the peptidyl carbonyl oxygen of Glu 148 to the amine group at N6 and another interaction from the indole nitrogen N1, to the peptidyl amine group of Ile 150 (Figure 2.9B). These interactions are observed in the afRio1-ATP complex (*24*). A hydrogen bond is seen between the 7-cyano group to a water molecule contacting the carbonyl side chain of Glu 120 and peptidyl amine group of Asp 212 (Figure 2.9B). This water molecule is seen in all available structures of afRio1 and in the afRio1-ATP complex this water mediates a hydrogen bond between Glu 120 and the **γ**-phosphate (*24*). The cyano group present in toyocamycin fits well in the

available pocket near the N7 position and hydrogen bonding to afRio1 via Glu 120, providing efficient binding to the protein. The ribose 3' OH group participates in two water-mediated hydrogen bond interactions; one with the side chain of Glu 162 and another water-mediated hydrogen bond with the peptidyl carbonyl oxygen of Tyr 200, an interaction not observed in the afRio1-ATP complex (Figure 2.9B). The ribose 2' OH group participates in a water-mediated hydrogen bond with the peptidyl carbonyl oxygens of Ile 55 and Ala 157, these interactions are not observed in the afRio1-ATP complex. The 5' OH has a direct hydrogen bond with Asp 212 as well as a water-mediated hydrogen bond with the peptidyl carbonyl oxygen of Tyr 200 (Figure 2.9B). In the afRio1-ATP complex the Asp 212 participates in one direct hydrogen bond to the γ-phosphate and two Mn^{2+}-coordinated interactions with the α- and β-phosphates of ATP (*24*). Despite the lack of the triphosphate moiety, toyocamycin is able to make a direct hydrogen bond with the metal binding residue, Asp 212.

2.5 Discussion

Despite the absence of the triphosphate moiety in toyocamycin and therefore the disruption of hydrogen and metal mediating interactions between the phosphates and afRio1, toyocamycin binds to afRio1 with a higher affinity than the native ligand, ATP/Mg^{2+}. In addition afRio1 shares eleven hydrogen bonds with ATP whereas toyocamycin has only ten hydrogen bonds to afRio1, which is equivalent to about 0.5-1.5 kcal/mol loss in energy for the toyocamycin complex. 7-methylinosine, on the other hand, showed the least stabilization as expected due to the absence of a hydrogen donor at C6 position necessary for optimal binding to afRio1. In addition,

repulsion between the carbonyl groups at the C6 position and the peptide bond of Glu 148 could be a contributing factor to the weaker binding observed with 7-methylinosine since this distance is about 3.1 Å.

In an effort to characterize the interactions of afRio1 and toyocamycin, we present a complex structure that also serves as the first structure of toyocamycin bound to a ribosome-processing factor, Rio1. Previous studies have indicated toyocamycin inhibits the maturation of 18S rRNA but there are no reports in literature that identify a binding target for toyocamycin during ribosome biogenesis. This crystal structure and the high binding affinity of toyocamycin to afRio1 provides a potential tool that can be used to identify Rio1 substrates which have yet to be determined. There are several reports on the use of radiolabeled ATP analogs such as N6-benzyl ATP as a tool for identifying specific kinase substrates (*78*). The use of ATP analogs promotes specificity for the kinase to be studied since other kinase are not likely to accept the analog (*78*). Using the same approach, toyocamycin 5'-triphosphate [γ-32P] can be introduced to cellular extracts from a system that has overexpressed Rio1 kinase. Proteins that are radiolabeled can be identified by mass spectrometry methods. This approach ensures that the overexpressed Rio1 is responsible for the transfer of γ-32 phosphate to the substrate and specificity of toyocamycin for Rio1 limits to some extent the possibility that other kinases are responsible for phosphorylation. The use of radiolabeled ATP analogs has been successful in many systems including in the identification of v-Src. ERK2 and CDK1 substrates (*78-80*). Some drawbacks of this approach in our system include the possibility of toyocamycin 5'-triphosphate binding to other kinases in the cellular

extract, not knowing the binding affinity of toyocamycin-5'triphophate for Rio1 and whether toyocamycin-5'triphophate has a higher affinity than ATP. The assumption is that toyocamyin binds to afRio1 with a high affinity, which does not necessarily mean that the triphosphate version would have the same binding affinity.

An alternative approach of using toyocamycin as a tool to identify substrates is based on a recent report that identified the translational elongation factor EEF1D as a substrate for casein kinase 2 (*81*). This functional proteomic strategy would require treating cells with toyocamyin in the presence of ATP [γ-32P] and identifying proteins with diminished phosphorylation by employing 2-D electrophoresis and mass spectrometry. Proteins with diminished phosphorylation can be determined by comparing the 2-D gel results to those of a control reaction (cells that don't have toyocamycin present). In order to validate identified proteins as direct substrates of Rio1, several approaches can be employed including: a) immuno-kinase assays that would require tagging the substrate and using antibodies to immuno-precipitate the substrate from cellular extracts followed by kinase reactions using Rio1, b) knockdown studies of Rio1 followed by determination of diminished phosphorylation of the substrates, c) designing inhibitor resistant mutants of Rio1 to evaluate the ability of these mutants to restore substrate phosphorylation in the presence of toyocmaycin.

2.6 Summary

This chapter identifies toyocamycin as a small molecule that binds to afRio1 in the ATP binding pocket and with a higher affinity than ATP. This provides the first

crystal structure of toyocamycin bound to a ribosome processing factor and a tool for identifying Rio1 substrates by employing strategies presented in the discussion. The next focus in this project was to determine whether toyocamycin inhibits the activity of afRio1. This will be discussed in the next chapter.

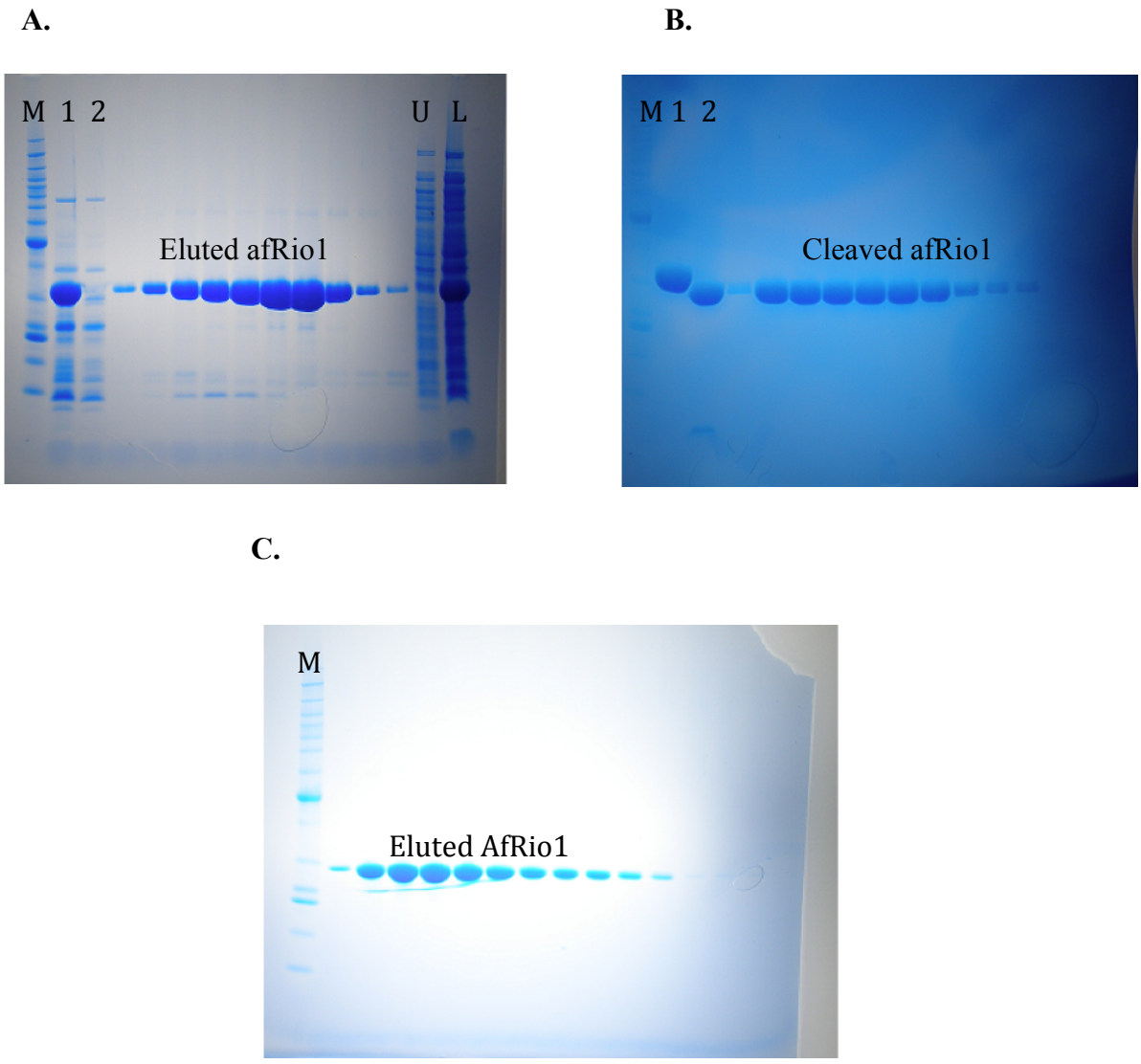


Figure 2.1: Preparative gels of afRio1 purification. Expected molecular weight is 30.09 kDa. Lane M in all gels is the protein ladder **A.** Gel after first HisTrap column purification, lane 1 and 2 are fractions that eluted with 10mM imidazole. Lane U is uninduced sample and Lane L is the lysate sample. **B.** Gel after the second HisTrap column purification, lane 1 is before histidine tag cleavage and lane 2 is after cleavage. **C.** Gel after size exclusion chromatography.

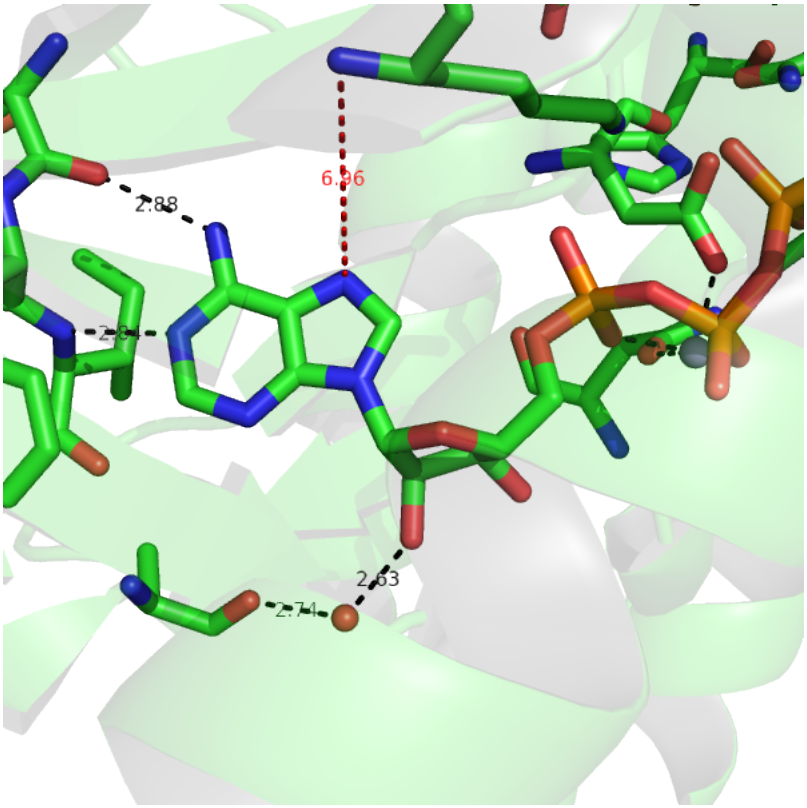


Figure 2.2: AfRio1-ATP active site. AfRio1-ATP complex showing an accessible pocket around the N7 position of the adenine moiety. The closest residue to N7 is 6.96 Å away as shown by the red dashed line. Black dashed lines are hydrogen bonds and the red sphere is water.

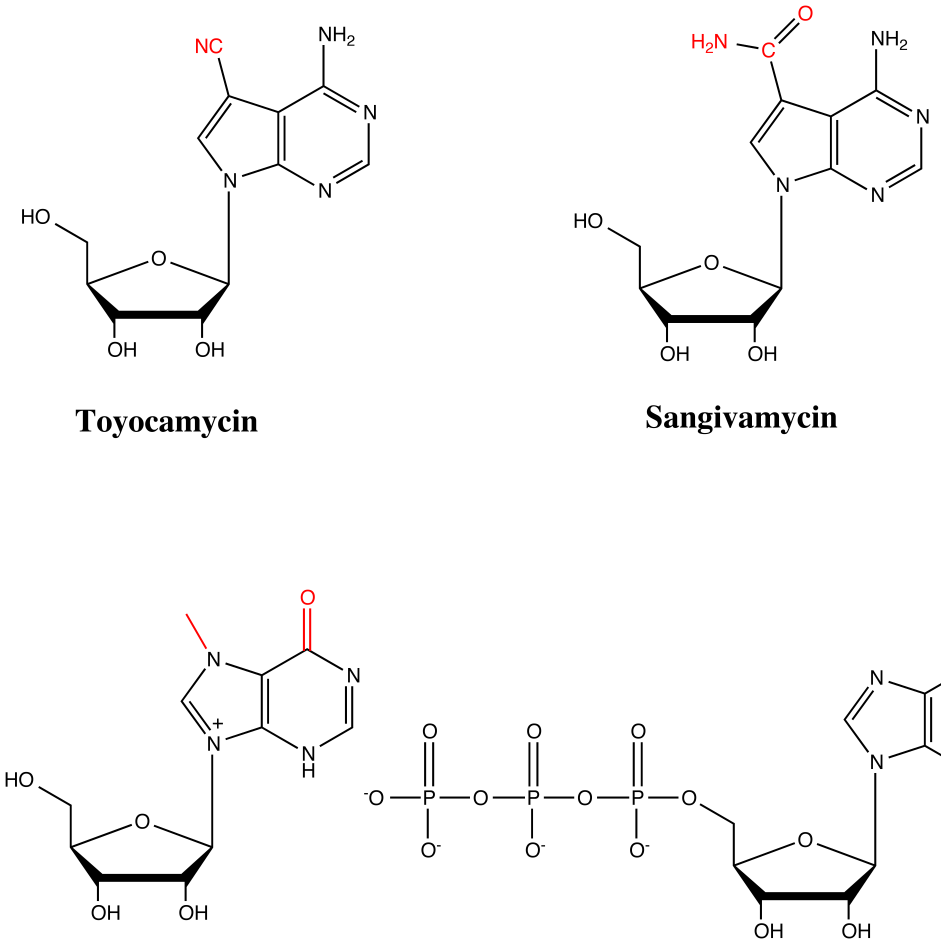
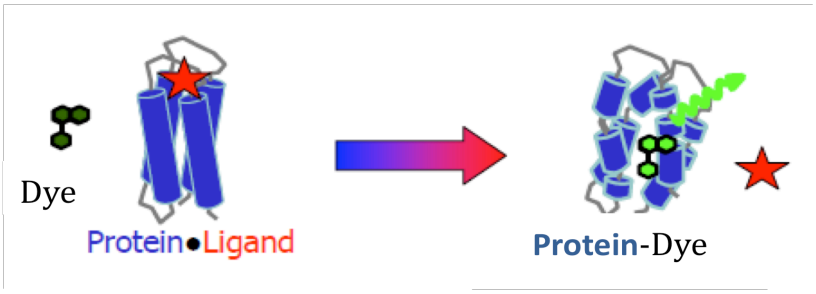


Figure 2.3: Small molecules. Chemical structure of small molecules tested for binding to afRio1.

A.



B.

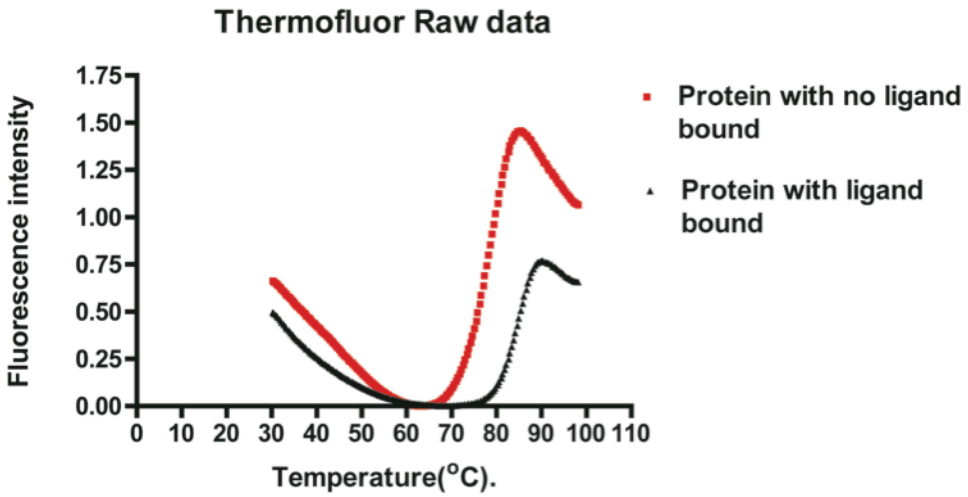
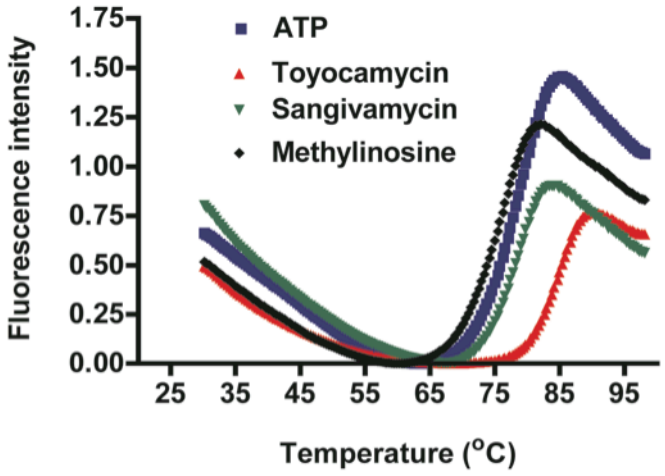


Figure 2.4: **Thermofluor assay**. **A.** Schematic showing the concept behind Thermofluor assay concept. **B.** Graph showing melting curves (red and black) generated after a thermofluor assay. A shift in melting temperature is observed with a stabilizing ligand (black curve).

A.



B.

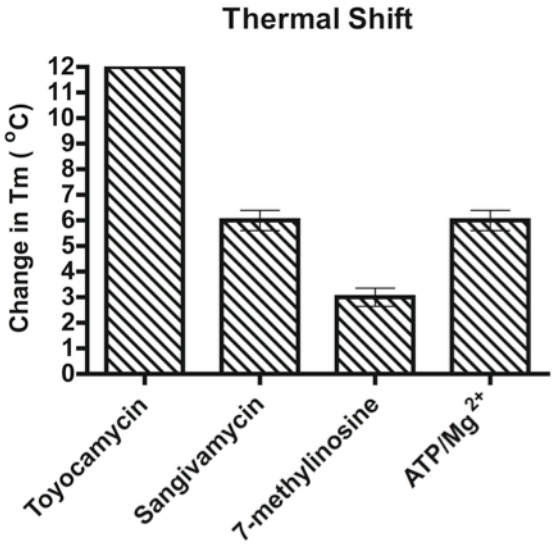
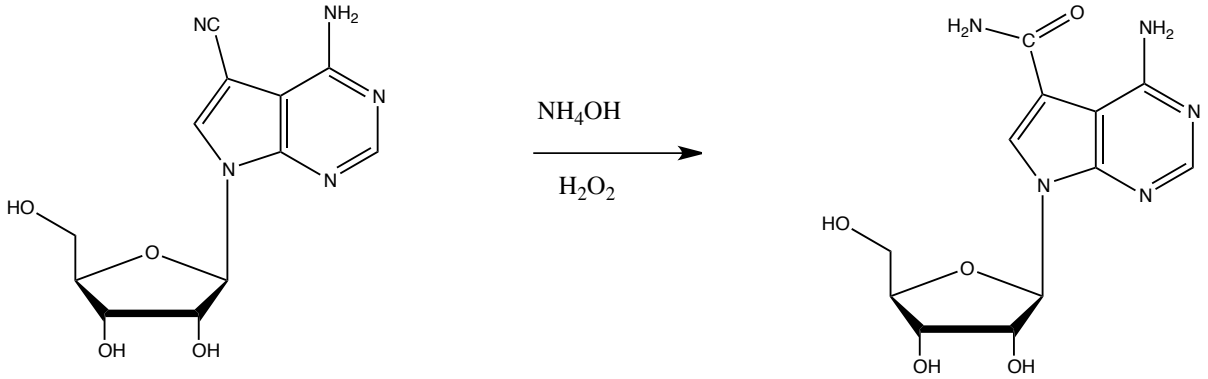


Figure 2.5: **Thermofluor assay data**. **A.** Melting curves of the small molecules tested. This is a representation of one replicate measurement. **B.** Graph showing change in melting temperature (T_m). Toyocamycin has the most change in T_m with a 12°C shift whereas 7-methylinosine has the least change in T_m, with a 3°C shift.



Toyocamycin Sangivamycin

Figure 2.6: Chemical conversion of toyocamycin to sangivamycin. The reaction proceeds by hydrolysis in the presence of the base NH_4OH.

A. **B.**

Figure 2.7: **Crystals of afRio1-toyocamycin complex**. **A**. Crystals of toyocamycin bound to afRio1 before optimization. **B**. Crystals of toyocamycin bound to afRio1after optimization.

Table 2:1 Scalepack logfile: AfRio1-toyocamycin complex.

Shell Lower limit	Shell Upper limit	Average Redundancy Per Shell
40.00	5.42	4.0
5.42	4.31	3.8
4.31	3.76	3.9
3.76	3.42	3.9
3.42	3.17	3.9
3.17	2.99	4.0
2.99	2.84	4.0
2.84	2.71	4.1
2.71	2.61	4.1
2.61	2.52	4.1
2.52	2.44	4.1
2.44	2.37	4.1
2.37	2.31	4.1
2.31	2.25	4.1
2.25	2.20	4.1
2.20	2.15	4.1
2.15	2.11	4.1
2.11	2.07	4.1
2.07	2.03	4.1
2.03	2.00	3.9
All hkl		4.0

Shell Lower limit	Shell Upper limit	I/Sigma in resolution shells: % of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
40.00	5.42	0.7	0.9	1.1	1.3	1.7	2.6	22.0	76.7	98.6
5.42	4.31	0.1	0.1	0.2	0.4	0.8	1.8	27.0	71.6	98.6
4.31	3.76	0.1	0.3	0.3	0.3	0.6	2.4	46.1	52.2	98.3
3.76	3.42	0.1	0.1	0.2	0.4	0.9	3.0	56.0	42.8	98.7
3.42	3.17	0.1	0.1	0.4	1.0	2.1	7.9	70.1	28.1	98.2
3.17	2.99	0.3	0.4	0.8	1.8	3.3	10.0	76.7	22.2	98.8
2.99	2.84	0.1	0.4	0.6	1.7	4.0	14.6	70.8	27.0	97.8
2.84	2.71	0.3	0.7	1.6	3.3	6.1	18.6	78.0	20.5	98.4
2.71	2.61	0.2	1.0	2.3	4.3	9.0	27.2	76.2	22.0	98.2
2.61	2.52	0.3	1.2	3.3	5.2	11.5	31.5	82.8	14.9	97.7
2.52	2.44	0.6	1.8	4.6	8.3	16.1	41.9	81.8	16.1	97.8
2.44	2.37	0.4	1.7	5.2	9.1	19.2	45.5	84.5	12.7	97.2
2.37	2.31	1.2	2.1	5.7	10.5	20.8	48.5	84.9	12.5	97.4
2.31	2.25	0.8	3.0	8.7	15.2	27.7	56.4	87.2	9.7	96.9
2.25	2.20	1.4	3.7	8.3	15.9	29.3	59.0	88.1	9.2	97.2
2.20	2.15	1.6	5.4	11.9	19.5	35.4	65.9	90.2	6.8	97.0
2.15	2.11	1.9	5.8	13.5	23.2	42.8	72.2	92.7	4.7	97.3
2.11	2.07	2.3	8.1	18.9	28.6	46.3	77.4	93.3	3.3	96.6
2.07	2.03	2.9	9.9	21.3	35.5	54.5	80.9	94.5	2.2	96.7
2.03	2.00	3.0	11.9	24.7	38.1	59.8	84.1	94.1	2.0	96.1
All hkl		0.9	2.9	6.6	11.1	19.5	37.4	74.7	23.0	97.7

Shell limit	Lower Angstrom	Upper	Average I	Average error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac
40.00	5.42	559.5	24.0	7.1	0.737	0.042	0.046	
5.42	4.31	601.9	27.2	9.0	0.836	0.046	0.052	
4.31	3.76	494.6	24.8	8.0	0.790	0.050	0.057	
3.76	3.42	349.1	18.1	6.8	0.914	0.056	0.064	
3.42	3.17	223.7	12.1	5.3	1.071	0.066	0.069	
3.17	2.99	162.3	9.0	4.3	1.159	0.072	0.072	
2.99	2.84	119.5	6.5	3.9	1.329	0.081	0.082	
2.84	2.71	98.3	5.6	3.5	1.318	0.086	0.082	
2.71	2.61	73.0	4.3	3.2	1.466	0.099	0.093	
2.61	2.52	63.1	4.1	3.2	1.471	0.112	0.107	
2.52	2.44	56.7	3.8	3.2	1.483	0.118	0.104	
2.44	2.37	52.0	3.8	3.2	1.423	0.125	0.110	
2.37	2.31	48.4	3.7	3.4	1.425	0.133	0.122	
2.31	2.25	42.7	3.7	3.4	1.403	0.150	0.126	
2.25	2.20	41.0	3.7	3.4	1.409	0.157	0.139	
2.20	2.15	34.0	3.6	3.4	1.312	0.179	0.164	
2.15	2.11	30.7	3.7	3.5	1.235	0.200	0.177	
2.11	2.07	25.5	3.6	3.5	1.152	0.224	0.195	
2.07	2.03	22.0	3.6	3.5	1.089	0.252	0.208	
2.03	2.00	20.2	3.7	3.7	1.085	0.282	0.243	
All reflections		158.5	8.7	4.5	1.209	0.069	0.057	

Table 2.2: Data Collection and Refinement statistics

PDB ID: 3REA	**Toyocamycin bound afRio1**
Data collection	
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	52.814, 72.607, 60.561
α, β, γ (°)	90, 90.172, 90
Molecules/Asym. unit	2
Wavelength (Å)	0.97919
Resolution (Å)	40-2.00
R_{sym} (last shell)	0.069 (0.282)
$I/\sigma I$	18.2 (5.5)
Completeness (%)	97.7 (96.1)
Redundancy	4.0 (3.9)
Refinement	
Resolution (Å)	34.9-2.00
R_{work}/R_{free} (%)	21.6/25.5
Residues	503
Waters	173
Mean B-factors (Å)^2 Residues /backbone	33.98/31.36
RMS deviations	
Bond lengths (Å)	0.008
Bond angles (°)	1.157
Ramachandran plot	
Favored	95.74%
Additional Allowed	4.26%
Disallowed	0.0%

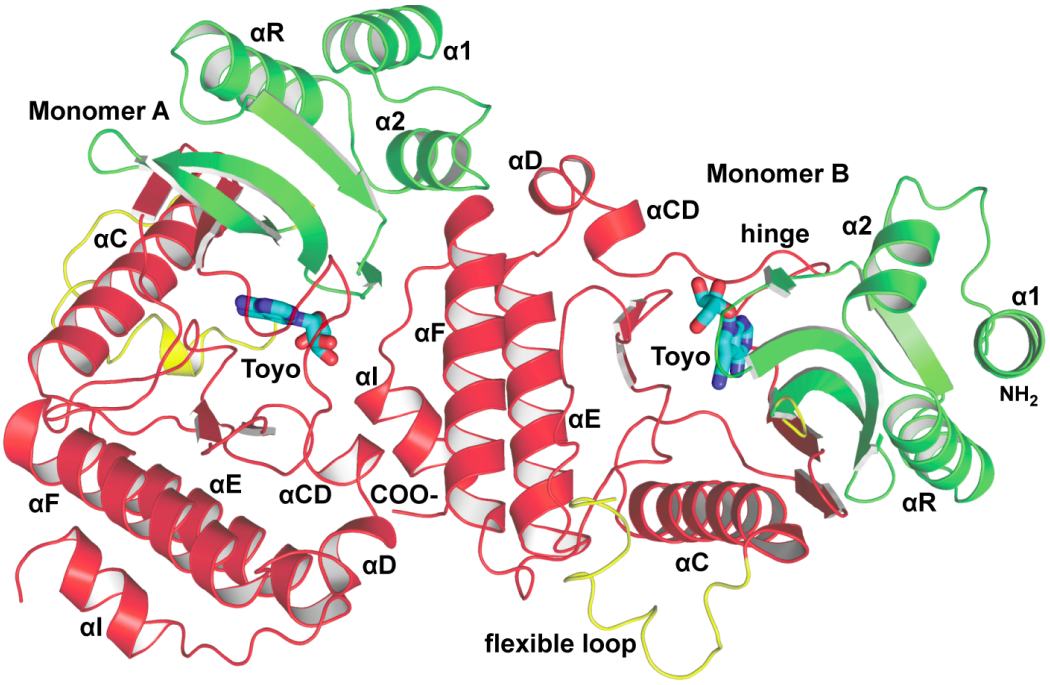


Figure 2.8: **AfRio1-toyocamycin overall structure with two molecules per asymmetric unit**. Toyocamycin is labeled as TOYO, the N-lobe is colored green and the C-lobe is red in both molecules.

A.

B.

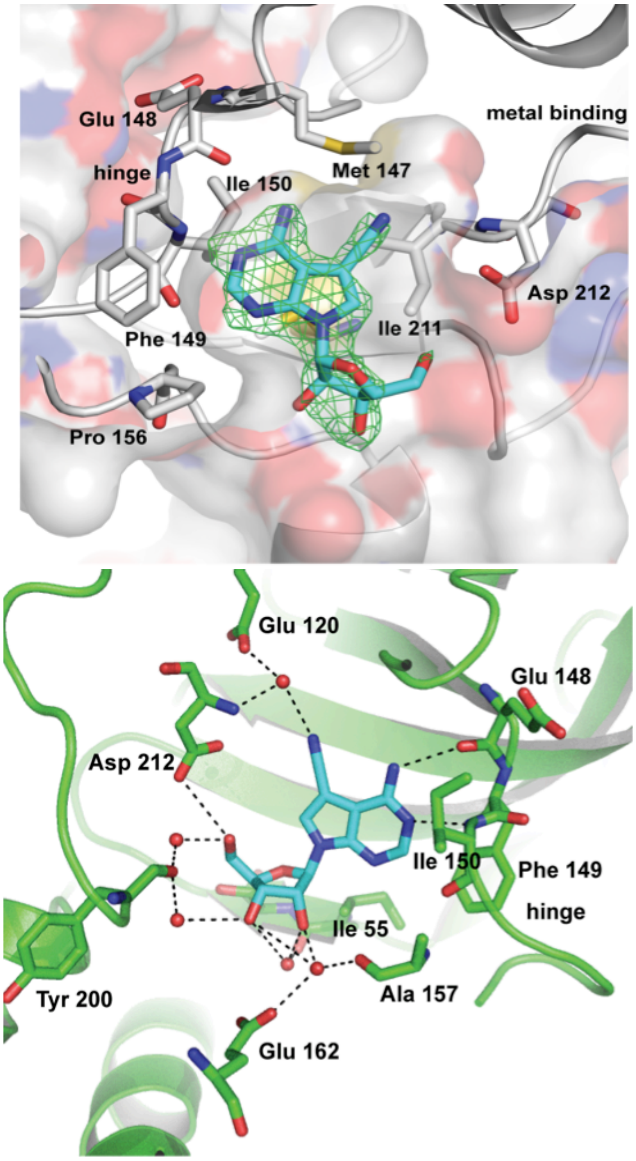


Figure 2.9: Toyocamycin binding site. **A**. Fo-Fc map at 2.5 σ showing electron density of toyocamycin in the ATP binding pocket. Residues involved in hydrophobic interactions with toyocamycin are labeled. **B**. Hydrogen bond interactions in afRio1-toyocamycin complex. Black dashed lines represent hydrogen bonds and red spheres are waters.

Chapter 3: Inhibition Studies and Oligomeric State Characterizations of *Archaeoglobus fulgidus* Rio1 (afRio1).

3.1 Overview

The discovery that toyocamycin binds to afRio1 with a higher affinity and stabilizes afRio1 more than ATP necessitated inhibition studies in order to determine whether toyocamycin was merely binding to afRio1 or causing inhibition. The first step was to determine the steady state kinetic parameters (K_m and V_{max}) of afRio1 for ATP, followed by characterization of the inhibition of afRio1's phosphorylation activity by toyocamycin. The crystal structure of afRio1-toyocamycin complex showed binding of toyocamycin to the ATP binding site leading to the hypothesis that toyocamycin was a competitive inhibitor. After analysis of kinetic parameters, a different type of inhibition model was used to characterize the mode of inhibition. This inhibition model required oligomeric state analysis by sedimentation equilibrium experiments as will be described in this chapter. In addition, mutational studies were carried out to determine the importance of a conserved residue, Tyr 200 that is located on the interacting interface of oligomers identified in the presence of ATP and toyocamycin.

3.2 Experimental Procedures:

3.2 a Steady State Kinetics studies

In order to determine the kinetic parameters (K_m and V_{max}) of afRio1 for ATP, 0.5 µg of purified afRio1 and 10 µg of myelin basic protein (MBP), a general kinase substrate that is phosphorylated by afRio1, was added to a 30 µl reaction volume in a buffer containing 50 mM Tris pH 8.0, 50 mM NaCl, 2 mM $MgCl_2$ and 0.2% β-mercaptoethanol. In order to monitor the phosphorylation activity, γ-^{32}P labeled ATP was used and the amount added varied from 1.0 x10^{-6} M to 7.9 x10^{-9} M. Each reaction was incubated at 37°C and stopped at specific time points (0, 1, 2, 4, 6, 8, 10, 15, 20, 30 and 40 minutes) by adding 8 µl of sodium dodecyl sulfate (SDS) loading dye and heating the reaction at 95°C for 5 minutes in order to denature the protein. All reactions were carried out in triplicates and the reaction mixtures were separated by running NuPAGE®4-12% Bis-Tris denaturing gels (Invitrogen) at 200V for 30 minutes. The gels were dried, exposed onto a phosphoimager screen for 15 hours and quantified using ImageQuant software (Molecular Dimensions) to determine moles of γ-^{32}phosphate that were incorporated into MBP. After quantifying, further analysis was carried out using GraphPad Prism. Initial velocity (V_o) was determined by fitting raw data to a linear regression plot of moles of γ-^{32} phosphate incorporated as a function of time in minutes (Figure 3.1). A plot of V_o as a function of ATP concentration [S] was fit to a nonlinear regression curve using the Michaelis-Menten equation, $V_o = (V_{max}[S])/(K_m + [S])$ and from this plot the V_{max} and K_m were determined (Figure 3.2). The same protocol was carried out for inhibition studies with

the addition of toyocamycin to the reaction mixtures at three different concentrations, 20 nM, 40 nM and 60 nM. The concentrations of toyocamycin used in these experiments were based on the dissociation constant (K_d) that had previously been determined by fluorescence to be 40 nM. The two concentrations at 60 nM and 20 nM are simply above and below the K_d.

3.2b Substrate inhibition by high concentrations of ATP

Inhibition studies also showed that afRio1 phosphorylation activity is inhibited by high concentrations of ATP. In an effort to determine a range of ATP concentrations to use in the inhibition studies, an experiment was carried out that varied the range of γ-^{32}P labeled ATP concentrations between 1.0×10^{-4} M and 1.0×10^{-8} M. In this experiment, 0.5 μg of purified afRio1 and 10 μg of MBP were added to a 30 μl reaction volume in a buffer containing 50 mM Tris pH 8.0, 50 mM NaCl, 2 mM $MgCl_2$ and 0.2% β-mercaptoethanol. The reactions were incubated at 37°C for 45 minutes and the components were separated by running a gel that was dried and exposed on a phosphoimager. Moles of γ-^{32}phosphate incorporated into MBP were quantified using ImageQuant and the experiment showed inhibition of afRio1 activity with ATP concentrations above 1.0×10^{-6} M (Figure 3.3A). A substrate inhibition plot was fit using nonlinear regression to the equation $V_o = V_{max}[S]/((K_m + [S])(1+[S]/K_i))$ in GraphPad Prism 5.0 (Figure 3.3B) (*82, 83*). From this fit, the parameters were determined as V_{max} = 3.34 pmol/sec (± 0.07), K_m = 0.76 μM (± 0.2) and K_i = 7.15 μM (±2.28)

3.2 c Sedimentation equilibrium

All protein samples were dialyzed against 50 mM Tris pH 8.0, 50 mM NaCl, 2 mM $MgCl_2$ and 0.2% β-mercaptoethanol. The samples were prepared in three different protein concentrations of 0.5 µg, 0.25 µg and 0.125 µg each in a final volume of 115 µl. Both phosphorylated and unphosphorylated forms of afRio1 in the presence of ATP/Mg^{2+}, toyocamycin, ADP and adenosine were analyzed by this method. In order to ensure that there was no absorbance interference during the experiment from toyocamycin, ATP, ADP and adenosine, the lowest concentration that had no absorbance at 280 nm was $1x10^{-5}$ M for all four compounds, within saturation levels. The experiments were carried out at 20°C in a Beckman Coulter Ultracentrifuge and data was collected at 280 nm using an optical absorption detection and six channel Epon centerpiece with a 12 mm path length. The sample volume was 115 µl for the samples and 120 µl for the buffer references. Data were collected for all samples were three speeds; 18000, 22000 and 26,000 rpm. Raw data was analyzed using the software, WinNonlin (*84*).

3.2d Site-directed Mutagenesis

Tyr 200 was mutated to Asp using QuikChange™ Lightning kit (Agilent Technologies) according to the manufacturer's instructions. Mutagenic primers used were

5'GGTGCATGCGGATCTGAGCGAAGATAACATTATGTATATTGATAAA 3' as the sense primer and

5'TTTTATCAATATACATAATGTTATCTTCGCTCAGATCCGCATGCACC 3' as

the antisense primer. Protein expression and purification of afRio1Y200D followed the same protocol as the wild-type afRio1 described in chapter 2 with only one modification. After lysis, the cells were heated to 68°C instead of 75°C for 15 minutes to denature *E.* coli proteins.

3.3 Results

3.3a Steady state kinetics

The Michaelis constant of afRio1 for ATP ($K_{m,ATP}$) was determined to be 0.43 μM (±0.11) and the maximum catalytic rate (V_{max}) was 2.55 pmol/sec (±0.29). The $K_{m,apparent}$ determined in the presence of a) 20 nM toyocamycin was 0.82 μM (±0.19), b) 40 nM toyocamycin was 0.23 μM (±0.03), and c) 60 nM toyocamycin was 0.59 μM (±0.02) (Table 3.1). The $V_{max,apparent}$ was 1.95 pmol/sec (±0.26) in the presence of 20 nM toyocamycin, 0.57 pmol/sec (±0.29) in the presence of 40 nM toyocamycin and 0.15 pmol/sec (±0.01) in the presence of 60 nM toyocamycin (Table 3.1). Crystallographic data suggested that toyocamycin was a competitive inhibitor since toyocamycin binds in the same site as ATP. However, the inhibition studies revealed that in the presence of inhibitor both V_{max} and K_m were changing, which is inconsistent with strict competitive inhibition that would show a change in K_m but not in V_{max}. This suggested that in addition to competitive inhibition, an uncompetitive mechanism was in play and a mixed inhibition model was assumed in order to account for both the occupancy of toyocamycin to the ATP binding site as well as a change in both K_m and V_{max}. Mixed inhibition occurs through allosterics with either the inhibitor binding to a different site from the substrate's (ATP) binding site or

through modulation of binding sites in oligomers. Since crystallographic data demonstrated that toyocamycin binds to the same site as ATP, the only other hypothesis was that inhibition occurred via modulation of binding sites in oligomers. This would imply that toyocamycin stabilizes an oligomeric form that is less catalytically active than the free form. However, before investigating the formation of oligomers, inhibition plots were fit to the mixed inhibition equation, $V_o=(V_{max}[S])/(\alpha K_m + \alpha'[S])$, where $\alpha = 1+([I]/K_i)$, $\alpha'=1+([I]/K_i')$, and K_i and K_i' are inhibition constants (Figure 3.4). The parameters α and α' were calculated using the equations: $K_{m,\ apparent} =\alpha K_m/\alpha'$, and $V_{max,\ apparent}= V_{max}/\alpha'$ respectively. The inhibition constants in the presence of a) 20 nM toyocamycin were 10.7 nM and 40 nM, b) in the presence of 40 nM toyocamycin were 26.7 nM and 11.4 nM, and c) in the presence of 60 nM toyocamycin were 3.8 nM and 5.5 nM (Table 3.1).

3.3b Substrate inhibition by high concentrations of ATP

The observed inhibition of afRio1 by high concentrations of ATP can be explained by substrate inhibition, a phenomenon that is observed in about 20% of known enzymes and occurs when more than one binding site is present (*82*). As substrate concentration increases, more sites are occupied and eventually the enzymes' activity is blocked (*82*). This is possible in oligomers that would have more than one binding site in a single entity. Therefore, inhibition of afRio1 by high concentration of ATP suggested ATP promoted oligomerization and stabilized a less catalytically active form of afRio1. In order to explain both mixed inhibition and

inhibition of afRio1 by high concentrations of ATP, sedimentation equilibrium experiments were carried out to determine Rio1 oligomeric states.

3.3c Sedimentation equilibrium

All data was initially fit to estimate average molecular weight using a model for single species:

$$c_r = c_{ro}\exp[\sigma(r^2/2 - r^2_0/2)]$$

where c_r is the concentration of the monomer at some radial position r, c_{ro} is the concentration at a reference position $r_{o,}$ and $\sigma = M(1-\upsilon\rho)\omega^2/ 2RT$ (M is the monomer molecular weight, υ is the partial specific volume of the protein, ρ is the buffer density, ω is the angular velocity, R is the gas constant and T is the temperature in Kelvin).

In almost every case except for phosphorylated afRio1 with no ligand bound, a single species model resulted in a poor fit. The best fit was determined by minimization of the $(\text{variance})^{1/2}$ and by examining the residuals. Data was therefore fit to various models; monomer, monomer-dimer, monomer-trimer, monomer-dimer-tetramer, and monomer-tetramer equilibrium. The results of the best-fit models are summarized in Table 3.2. The monomeric molecular weight of afRio1 is 30.1 kDa, the unphosphorylated afRio1 that has residual adenosine in the active site fit to an average molecular weight of 60.2 kDa (± 1.1) but the best fit was a monomer-tetramer model, efforts to fit this data to a dimer or monomer-dimer resulted in both poor residuals and $(\text{variance})^{1/2}$ (Figure 3.5A). When toyocamycin is bound to unphosphorylated afRio1, the average molecular weight is larger than a dimer,

71.8 kDa (±1.2) and the best-fit model is monomer-tetramer (Figure 3.5B). In the presence of ATP or ADP with Mg^{2+}, the average molecular weight was between monomer and dimer, 53.6 kDa (±1.1) for ATP and 50.8 kDa (± 1.1) for ADP, and the best-fit model for both was monomer-dimer (Figure 3.5C-D). However, when afRio1 is phosphorylated the average molecular weight is 34.1 kDa (± 1.0) and fits to a monomer model, suggesting that phosphorylation prevents oligomerization (Figure 3.5E). Interestingly when toyocamycin and ATP are added to this monomeric form of afRio1 (phosphorylated afRio1), oligomerization is observed with increased average molecular weights of 46.2 kDa (± 1.2) and 47.6 kDa (± 1.1) respectively and both fitting to a monomer-trimer model (Figure 3.5F-G). This data showed that in the presence of toyocamycin, ATP and ADP, the unphosphorylated afRio1 protein forms higher order oligomers and phosphorylation promotes monomerization. Addition of ATP and toyocamycin to the monomeric afRio1 resulted in an increased average molecular weight and oligomerization. Taken together the data suggested that ATP and toyocamycin promote oligomerization consistent with the proposed mixed inhibition model by toyocamycin as well as substrate inhibition of afRio1 by high concentrations of ATP.

The dissociation constants (K_d) for oligomerization were also calculated from the sedimentation experiments and are summarized in Table 3.2. By comparing the monomer-tetramer mixtures seen in solution, afRio1-toyocamycin has a lower K_d, 1.47×10^{-6} M, compared to afRio1 with residual adenosine that has a K_d of 6.11×10^{-6} M suggesting afRio1-toyocamycin has a tighter complex. In the presence of the two ligands that promote some extent of dimerization, ATP and ADP, the results showed

that afRio1-ATP has a tighter complex (K_d=3.12 x 10^{-6} M) compared to afRio1-ADP that has a K_d of 6.30 x 10^{-6} M.

3.4 Discussion

3.4a Understanding interacting interfaces in afRio1 oligomers.

Sedimentation equilibrium analysis showed that afRio1 forms oligomers depending on the ligand bound. To better understand possible interacting interfaces involved in oligomerization, crystal structure lattice packing were analyzed. Previously solved crystal structures of unphosphorylated afRio1 with residual adenosine in the active site and those bound to ATP/Mn^{2+} and ADP/Mn^{2+} were available for analysis (*24*). These structures were used along with afRio1-toyocamycin complex structure to compare packing interfaces and identify structural differences that may promote formation of oligomers. It was previously reported that unphosphorylated afRio1 with residual adenosine differed from afRio1 bound to ATP/Mn^{2+} in positioning of the flexible loop and in the conformation of N-lobe helices (*24*).

The PDBe PISA (protein interfaces, surfaces and assemblies) server was used to analyze the crystal structures' interfaces as well as calculate the buried surface area between the interfaces (*85, 86*). In the structure of afRio1-toyocamycin complex, the largest interface had a buried surface area of 836.5 $Å^2$ per molecule and is observed between two molecules found in the asymmetric unit. The ATP/Mn^{2+} and ADP/Mn^{2+} bound crystal structures each have four molecules per asymmetric unit, but employ

the same packing interfaces observed in the afRio1-toyocamycin complex, with two copies of a similar dimer per asymmetric unit. The largest interface for both afRio1-ATP (buried surface area 766.6 $Å^2$) and afRio1-ADP (buried surface area 768.4 $Å^2$) bound structures is similar to the largest interface in the afRio1-toyocamycin complex (Figure 3.6). This interface positions the helices of the C-lobe of one monomer, monomer B, in contact with both the N- and C-lobes of the second monomer, monomer A, almost completely occluding the active site (Figure 3.6). The N-lobe helix, α2, of monomer A and the C-lobe helix (αF) of monomer B interact via hydrogen bonds between between A:Glu 23, A:Glu 22 and B:Arg 237; A:Glu 22 and B:Lys 241; A:Lys 25 and B:Tyr 242; A:Glu 29 and B:Lys 168 (Figure 3.6). The C-lobe of monomer A interacts with the C-lobe of monomer B via hydrophobic contact between conserved Tyr 200 of monomer A and Leu 252, Lys 253 and Phe 248 from the terminal alpha helix of monomer B (Figure 3.7). In addition, there is a hydrogen bond between A:Glu 162 and B:Arg 233. Glu 162 also participates in a water-mediated hydrogen bond with the 3' OH of toyocamycin's ribose moiety (Figure 3.7). This interface is not observed in the crystal structure of afRio1 with residual adenosine. The largest interface observed in the adenosine bound structure has a surface area of 1023.3 $Å^2$ and instead positions the N-lobe of one monomer (monomer B) between the N- and C-lobes of a symmetry related monomer (monomer A (Figure 3.6). The interface is extensive and positions the first three N-terminal helices (α1, α2, αR) of monomer B in contact with monomer A. This interface does not result in occlusion of the active, and in fact, a second symmetry related molecule can positions its flexible loop that contains the autophosphorylation site (Ser 108), of

one monomer (monomer A) at the active site entrance to monomer B. Analysis of these interfaces reveal two different types of interaction surfaces and stabilization of two different types of dimers. The dimer interface observed in the crystal structure of afRio1-adenosine complex is different from the interface observed in the afRio1-toyocamycin and afRio1-ATP complex. The afRio1-toyocamycin complex has the same packing as the afRio1-ATP/Mn^{2+} and afRio1- ADP/Mn^{2+} complexes. This indicates that afRio1 exists in at least two distinct dimeric or oligomeric states, and that the afRio1-toyocamycin complex favors the same conformation observed in the afRio1-ATP complex. In addition, an alignment of afRio1-adenosine, afRio1-toyocamycin and afRio1-ATP dimers, reveals a good alignment between afRio1-toyocamycin and afRio1-ATP whereas afRio1-adenosine does not align well with the other two complexes, indicative of two different types of dimers (Figure 3.8). Despite the fact that toyocamycin is an adenosine analog and lacks the triphosphate moiety, the toyocamycin bound structure favors the ATP/Mn^{2+} bound conformation.

3.4b Disrupting interacting interfaces and oligomerization

Analysis of the afRio1-toyocamycin, afRio1-ATP/Mn^{2+} and afRio1-ADP/Mn^{2+} complexes show the involvement of one key residue, Tyr 200, in the interacting interfaces. This residue is either phenylalanine or tyrosine in all Rio kinases, and is part of the catalytic loop containing the catalytic Asp (D196 in afRio1) and Asn (N201 in afRio1) residues. In order to determine the importance of Tyr 200 in dimerization of afRio1 when toyocamycin or ATP is bound, Tyr 200 was mutated to aspartate (Y200D) in order to disrupt aromatic hydrophobic interactions between

Tyr 200 and Phe 248 from the adjacent molecule, followed by oligomeric state analysis.

Oligomeric state analysis of afRio1Y200D was determined using sedimentation equilibrium experiments in order to confirm whether this mutation was affecting oligomerization by disrupting interaction interfaces (Table 3.3). The experiments were carried out in the presence and absence of bound ligand. In the absence of a ligand, sedimentation equilibrium data for the Y200D mutant had an average molecular weight of 70.8 kDa (± 1.1) and the best fit was to monomer-tetramer model based on evaluation of residuals as well as minimization of $(\text{variance})^{1/2}$ (Table 3.3, Figure 3.9A). This indicates that oligomerization is not significantly affected by the mutation in the unbound mutant. However, when toyocamycin or ATP/Mg^{2+} is added, the average molecular weight drops to 48.2 kDa (± 1.1 kDa) and 46.8 kDa (± 1.1 kDa) respectively, with the best fit to a monomer-trimer model in both cases (Table 3.3, Figure 3.8 B-C). The average molecular weights of the Y200D mutant bound to toyocamycin and ATP/Mg^{2+} are lower than that for the wild-type complexes in both cases showing more contribution from the monomeric species in the mutant (Table 3.3). The wild-type bound to toyocamycin and to ATP/Mg^{2+} has an average molecular weight of 71.8 kDa (±1.2) and 53.6 kDa (±1.1) respectively. Therefore there's a significant change in comparison to the mutant average molecular weights indicating that oligomerization in the presence of these ligands is compromised by mutation of Tyr 200. In addition the wild-type in the presence of toyocamycin fits to a monomer-tetramer model yet with the mutant the best fit is to a monomer-trimer model.

3.4c Formation of different afRio1 oligomeric states

The dimeric afRio1 structures described in this chapter (adenosine, ATP and toyocamycin complexes) form open oligomers that have a second interface available for higher order oligomerization. In addition the interacting interfaces are heterologous, meaning the interfaces are formed by two different surfaces from interacting monomers. Therefore, tetramers observed in solution and expected tetramers could be formed from interaction between dimers (Figure 3.10). Analysis of the adenosine complex revealed that interaction between dimers that utilize the largest interaction interface would bury an additional 1055 $Å^2$ of surface area, and present the flexible loop containing the autophosphorylation site, Ser108, to the active site. Toyocamycin and ATP tetramers have additional buried surface areas of 881 $Å^2$ and 819 $Å^2$ respectively. Therefore, it is possible that dimer formation nucleates higher order oligomerization, and anything that stabilizes the dimer would promote formation of the tetramer. In this model, trimers could be formed by addition of a monomer to a dimer, and would be observed when the proportion of monomer is increased, as seen for phosphorylated afRio1 in the presence of ATP/Mg^{2+} and toyocamycin.

In understanding how afRio1 form higher order oligomers, it is also important to examine how afRio1 switches between different oligomeric forms. Toyocamycin is an adenosine analog and based on Thermofluor data toyocamycin binds with a higher affinity and stabilizes afRio1 more than ATP/Mg^{2+}. Despite these results afRio1-toyocamycin complex adopts the same conformation as the afRio1-ATP complex.

This is contradicts predictions that due to their very similar structures, toyocamycin and adenosine would adopt the same conformation and stabilize the same type of oligomers. Therefore, the hypothesis was that there must be a structural conformational change or switch that influences the choice of interacting interface conserved between the ATP and toyocamycin bound structures but is different in the adenosine bound structure. Upon closer inspection of the active sites of the afRio1 structures with adenosine, toyocamycin and ATP bound, a difference in positioning of the metal binding loop containing the metal binding residue, Asp 212 was observed (Figure 3.11). Specifically, in the adenosine bound structure, the backbone carbonyl oxygen bond of Ile 211 points into the binding pocket towards the nucleoside, while in the ATP and toyocamycin bound structures, there is a peptide flip that positions the same carbonyl oxygen in the opposite direction. This results in a 3.8 Å shift in the carbonyl oxygen position producing a shift in the metal binding loop containing residues Ile 211 to Gln 215. Subsequently the Cα of Gln 215, a residue that is conserved in all RIO kinases, is shifted by 2.1 Å and the side chain is within hydrogen bonding distance to Asp 212 carboxylate oxygen. The flexible loop of afRio1 bound to toyocamycin and ATP is shifted away away from the active site and a shift in the catalytic loop is also observed (Figure 3.11).

3.5 Summary

Despite crystallographic studies that show binding of toyocamycin to the ATP binding site that suggests strict competitive, steady state kinetics revealed mixed inhibition as the mode of inhibition. In addition to inhibition by toyocamycin,

substrate inhibition of afRio1 by high concentrations of ATP is observed. Both mixed inhibition and substrate inhibition are supported by formation of oligomers in the presence of toyocamycin and ATP. These oligomers are less catalytically active than the afRio1 free form. Analysis of interacting interfaces and formation of oligomers reveals the stabilization of two different types of dimers as a result of conformational switches. AfRio1-ATP and afRio1-toyocamycin complexes form the same type of dimer whereas the afRio1-adenosine forms a different type of dimer based on alignment and interface analysis. These dimers are open, heterologous and are therefore capable of forming higher order oligomers. This validates the monomer-trimer and monomer-tetramer models that are observed in solution. The question remained whether afRio1 studies are applicable to eukaryotic Rio1. The answer to this question is part of the discussion in the next chapter.

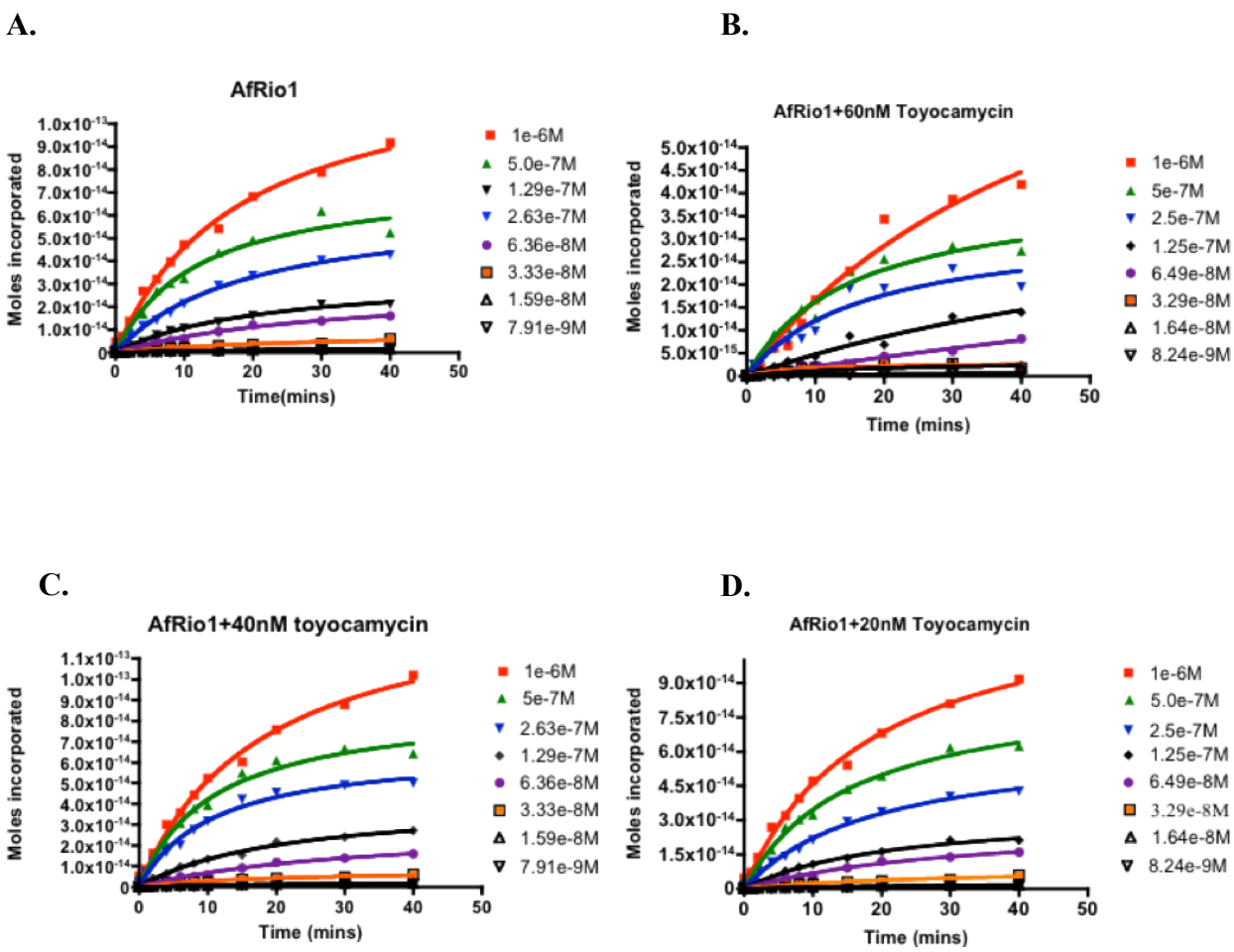


Figure 3.1: **Plots of moles of γ-^{32}phosphate incorporated as a function of time.** The plots represent the average moles of γ-^{32}phosphate since reactions were done in triplicates. Initial velocity (V_o) was determined as the linear slope. **A.** Plot of afRio1 with no toyocamycin added to the reaction mixtures. **B.** Plot of afRio1 with 60 nM toyocamycin added to reaction mixtures. **C.** Plot of afRio1 with 40 nM toyocamycin added to reaction mixtures. **D.** Plot of afRio1 with 20 nM toyocamycin added to reaction mixtures.

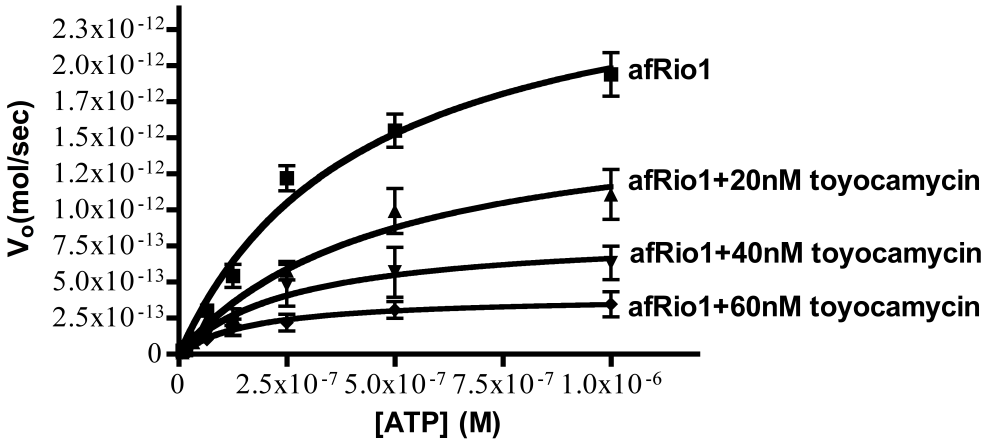
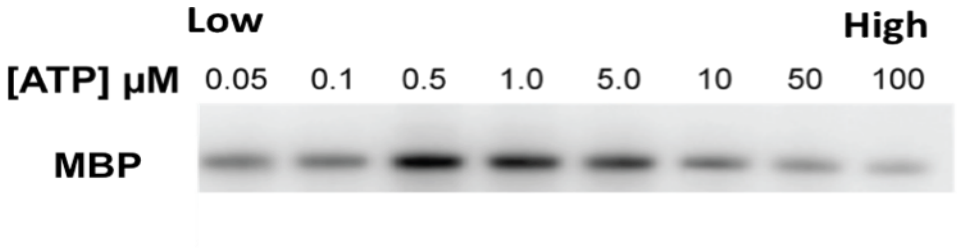


Figure 3.2: **Inhibition of afRio1 phosphorylation activity by toyocamycin**. Plot of initial velocity (V_o) as a function of ATP concentration. This plot was generating by fitting the plot of moles γ-^{32}phosphate incorporated as a function of time to the Michealis-Menten equation, $V_o = (V_{max}[S])/(K_m + [S])$. From the plots, the V_{max}, K_m, $V_{max,apparent}$ and $K_{m,apparent}$ were determined. Errors bars indicate standard deviations.

A.



B.

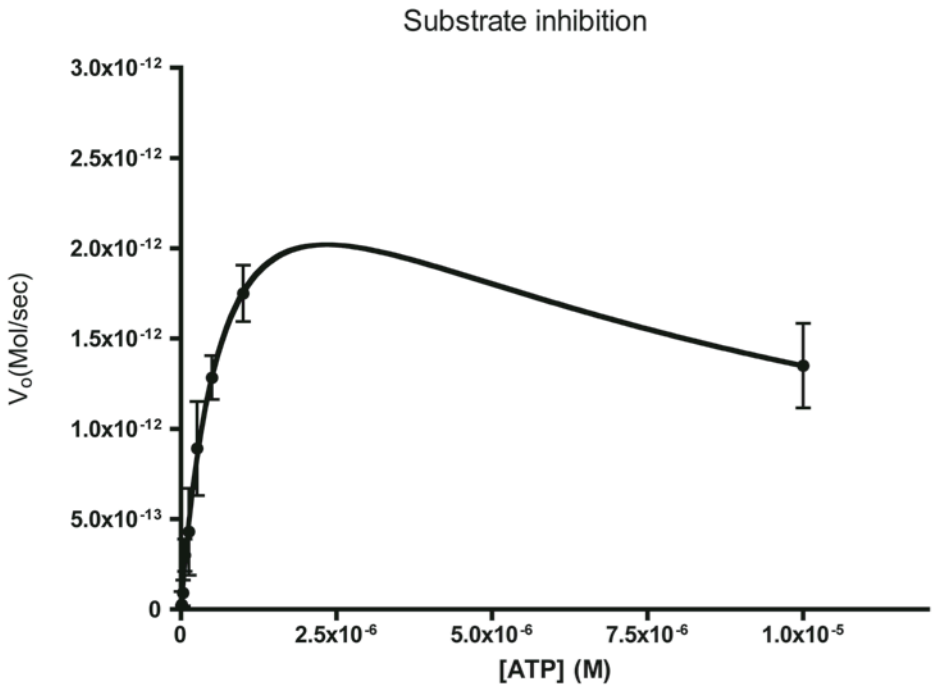


Figure 3.3: Substrate inhibition. **A.** Gel showing inhibition of afRio1 phosphorylation activity by high concentrations of ATP above 1 µM. **B.** Graph fit to $V_o=V_{max}[S]/((K_m + [S])(1+[S]/K_i))$ showing substrate inhibition.

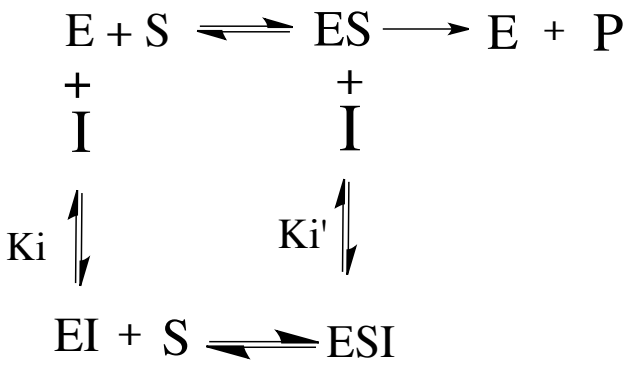


Figure 3.4: Schematic of mixed inhibition.

Table 3.1: Steady state parameters for afRio1 in the presence and absence of toyocamycin.

	afRio1	**+20 nM Toyocamycin**	**+40 nM Toyocamycin**	**+60 nM Toyocamycin**
K_m (μM)	0.43 (±0.11)	0.82 (±0.19)	0.23 (±0.03)	0.59 (±0.02)
V_{max} (pmol/sec)	2.55 (±0.29)	1.95 (±0.26)	0.57 (±0.29)	0.15 (±0.39)
Ki (nM)		10.7	26.7	4.6
Ki' (nM)		40.0	11.4	8.5

Table 3.2: Sedimentation equilibrium data for afRio1 bound to various ligands. AfRio1 monomer molecular weight is 30.1 kDa

Sample	**Averaged MW (kDa)**	**Best Fit Model**	**(Variance)^{½}**	**Estimated K_d (M)**
afRio1	60.2 ± 1.1	Monomer-tetramer	8.0×10^{-3}	6.11×10^{-6}
afRio1+ toyocamycin	71.8 ± 1.2	Monomer-tetramer	1.8×10^{-2}	1.47×10^{-6}
afRio1+ATP	53.6 ± 1.1	Monomer-dimer	2.8×10^{-2}	3.12×10^{-6}
afRio1+ADP	50.8 ± 1.1	Monomer-dimer	2.4×10^{-2}	6.30×10^{-6}
Phosphorylated afRio1	34.1 ± 1.0	Monomer	3.1×10^{-2}	N/A
Phosphorylated afRio1+ATP	47.6 ± 1.1	Monomer-trimer	8.8×10^{-3}	6.92×10^{-6}
Phosphorylated afRio1+toyocamycin	46.2 ± 1.2	Monomer-trimer	1.3×10^{-2}	6.38×10^{-6}

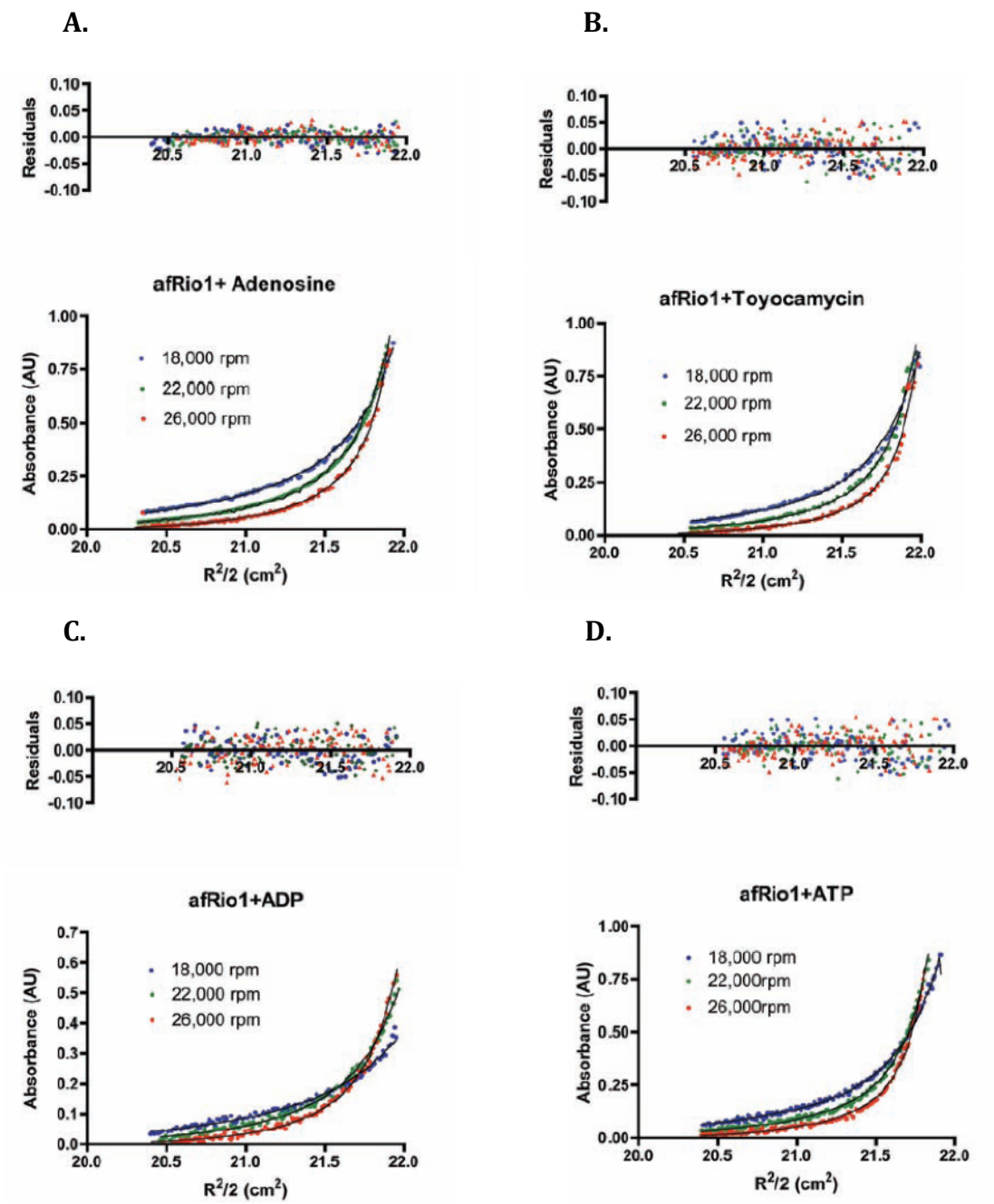


Figure 3.5: Sedimentation equilibrium data collected at three speeds 18,000 rpm (blue), 22,000 rpm (green) and 26,000rpm (red). The fitted data was analyzed at three speeds and at the same concentration (0.25 μg/μl). **A**. AfRio1 with residual adenosine. **B**. AfRio1 with toyocamycin bound. **C**. AfRio1 with ADP bound. **D**. AfRio1 with ATP bound.

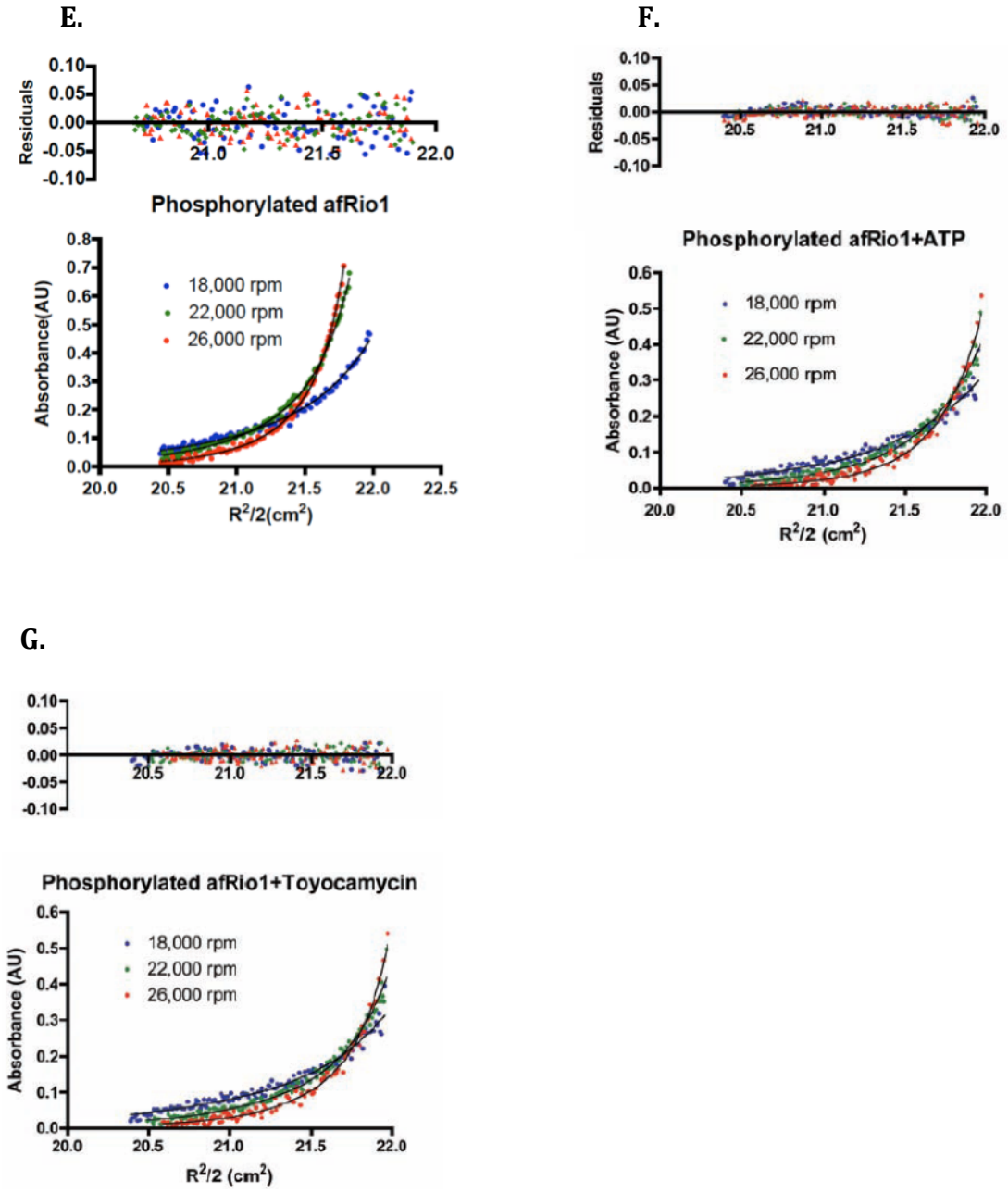


Figure 3.5: Sedimentation equilibrium data collected at three speeds 18,000 rpm (blue), 22,000 rpm (green) and 26,000rpm (red). The fitted data was analyzed at three speeds and at the same concentration (0.25 μg/μl). Each plot is accompanied with the respective residuals. **E.** Phosphorylated afRio1. **F.** Phosphorylated afRio1 with ATP bound. **G.** Phosphorylated afRio1 with toyocamycin bound.

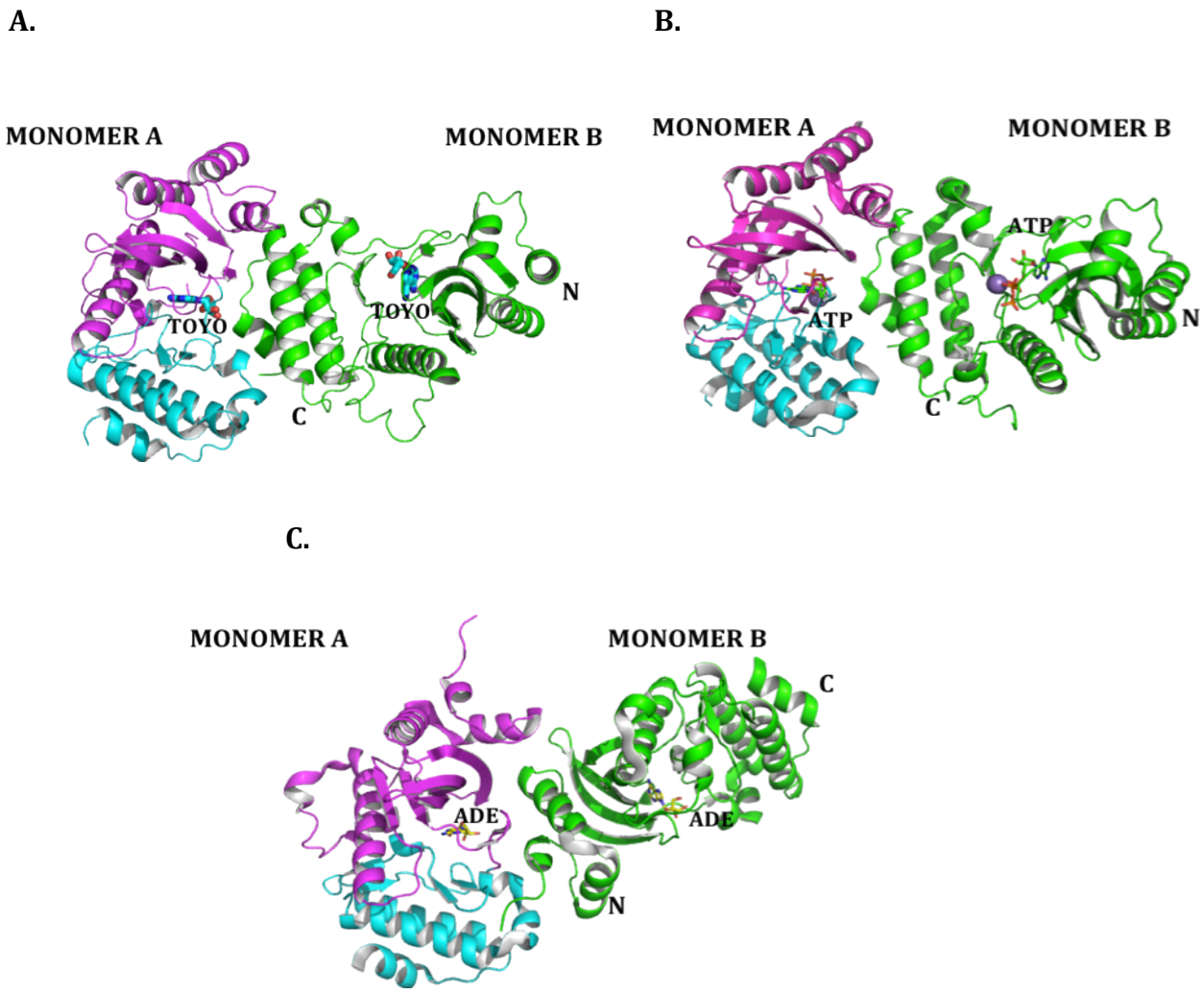


Figure 3.6: Dimer interfaces. The N-lobe of monomer A is in magenta and the C-lobe is in cyan whereas monomer B is in green. **A.** AfRio1-toyocamycin dimer with the C-lobe of monomer B interacting with N and C-lobes of monomer B. **B.** AfRio1-ATP dimer with the C-lobe of monomer B interacting with N and C-lobes of monomer B. **C.** AfRio1-adenosine with the N-lobe of monomer B interacting between the N and C-lobes of monomer A.

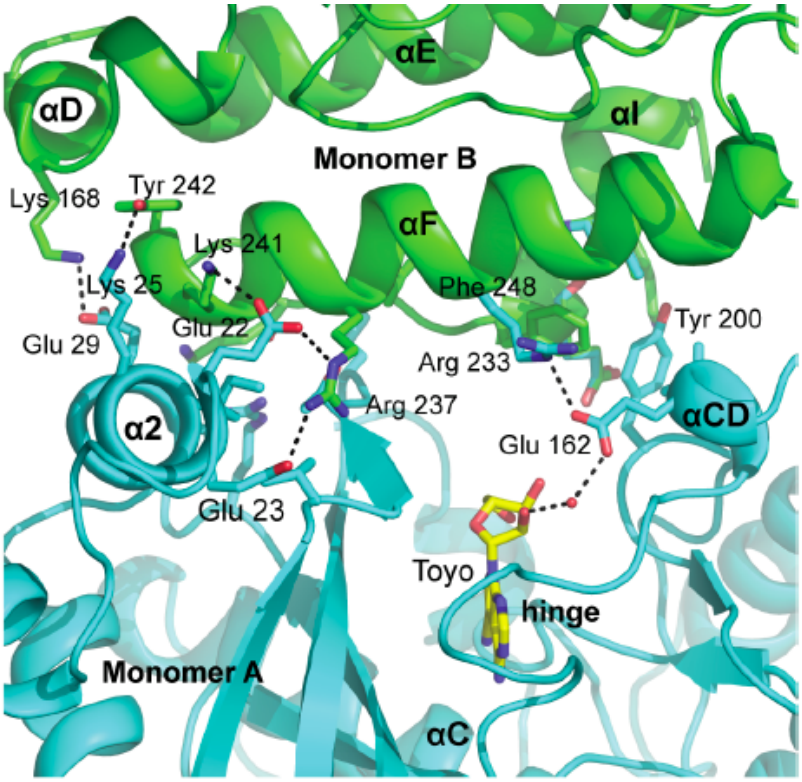


Figure 3.7: AfRio1-toyocamycin dimer interface. AfRio1-toyocamycin dimer interface showing hydrogen bonds (dashed black lines) between the C-lobe of monomer B and the N- and the C-lobes of monomer A. The dimer interface also indicates the occlusion of the active site and hydrophobic contact between conserved Tyr 200 of monomer A and Phe 248 of monomer B.

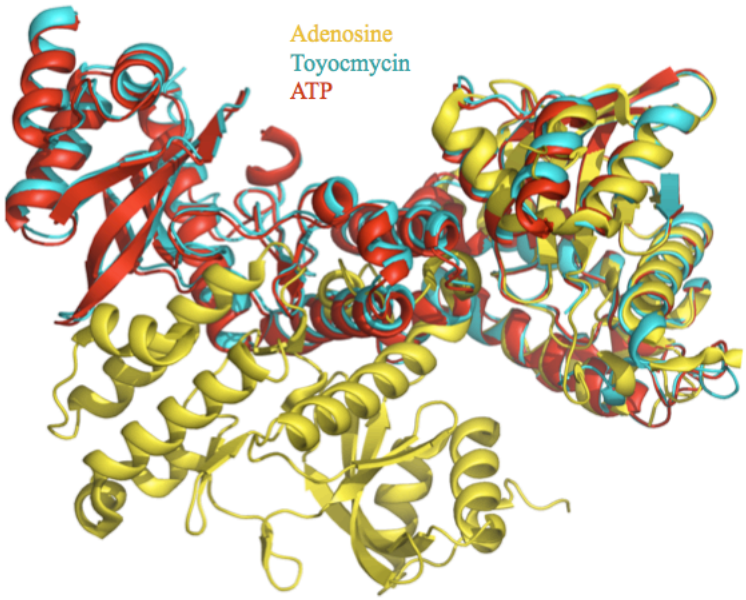


Figure 3.8: **Alignment of afRio1 dimers.** AfRio1-ATP is in red, afRio1-toyocamycin is in cyan and afRio1-adenosine is in yellow. AfRio1 bound to ATP and toyocamycin align well whereas afRio1 with adenosine does not align well with the other two dimers. This is evident of two different types of dimers, the ATP and toyocamycin stabilized dimer and the adenosine stabilized dimer.

Table 3.3: Sedimentation equilibrium data for afRio1 Y200D mutant

Sample	Averaged MW (kDa)	Best Fit Model	(Variance) ^{½}	Estimated K_d (M)
afRio1 Y200D	70.6 ± 1.1	Monomer-tetramer	9.2 x 10^{-3}	1.00 x 10^{-7}
afRio1 Y200D +toyocamycin	48.2 ± 1.1	Monomer-trimer	9.3 x 10^{-3}	1.20 x 10^{-5}
afRio1Y200D +ATP	46.8 ± 1.1	Monomer-trimer	9.4 x 10^{-3}	7.32 x 10^{-6}

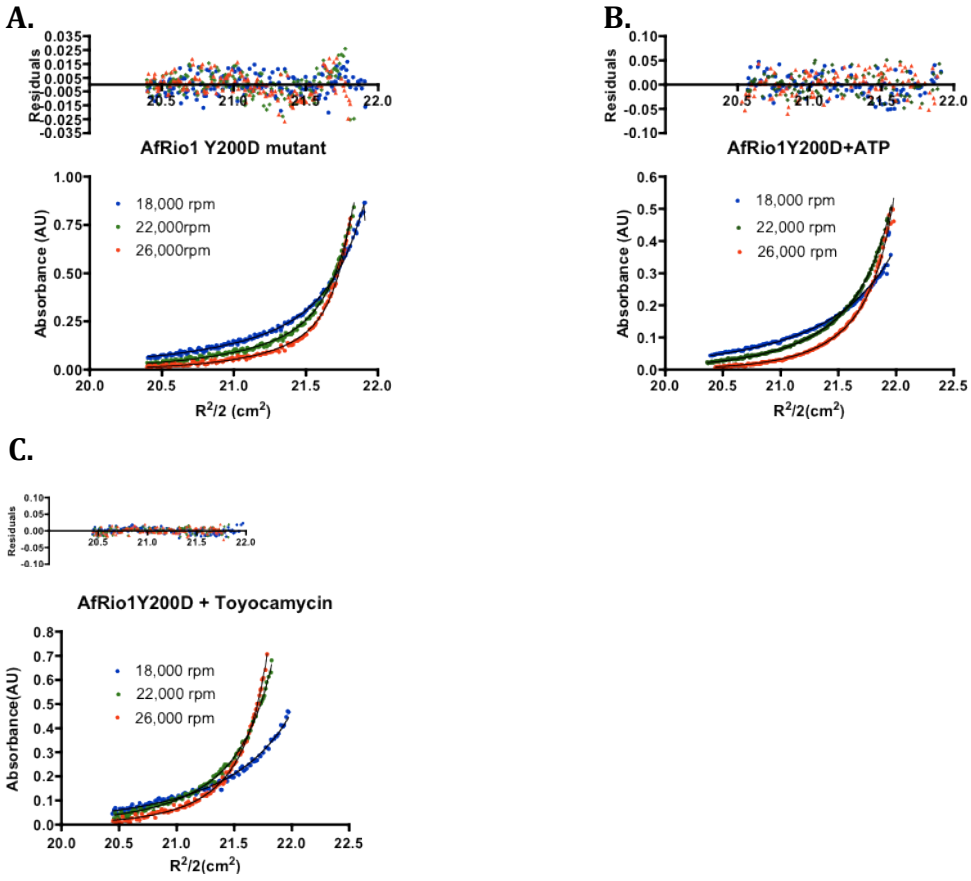


Figure 3.9: Sedimentation Equilibrium data for afRio1 Y200D mutant. A. Y200D mutant with no ligand bound. **B.** Y200D with ATP bound. **C.** Y200D with toyocamycin bound. Each plot is accompanied by the respective residuals.

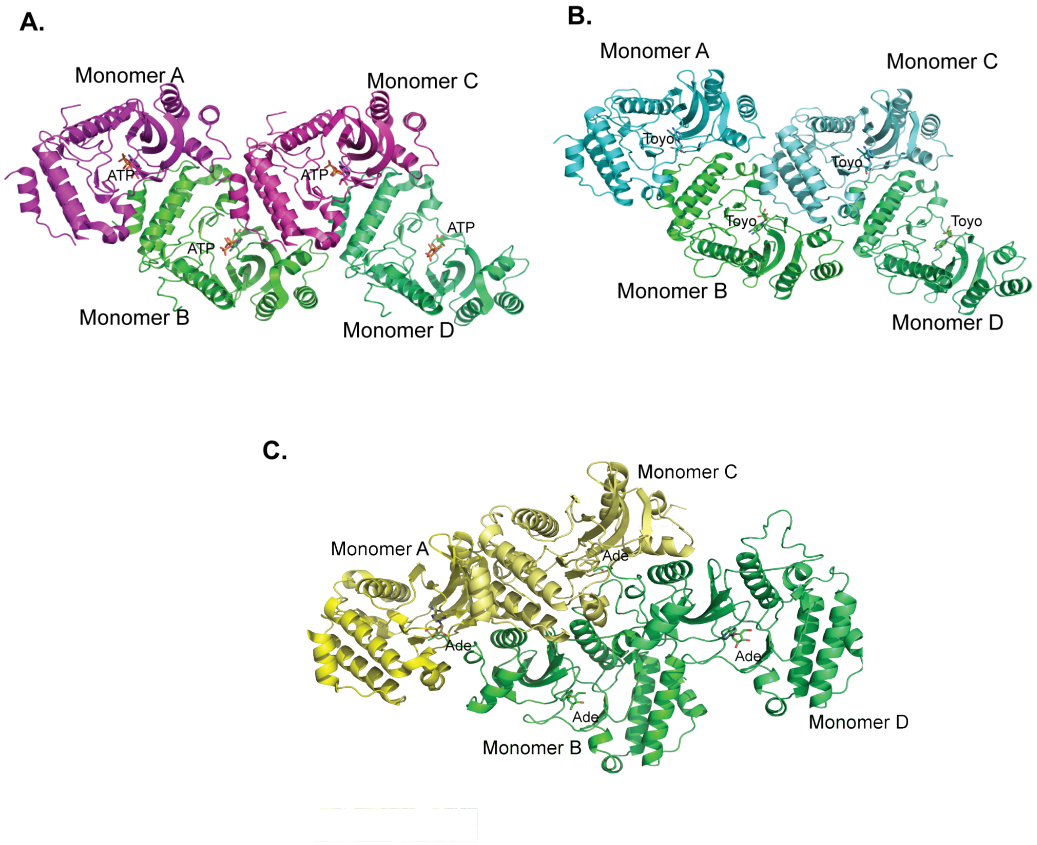


Figure 3.10: AfRio1 tetramers. ATP and toyocamycin stabilize the same type of tetramer, different from the afRio1-adenosine tetramer. **A.** AfRio1-ATP tetramer, each monomer has ATP bound. **B.** AfRio1-toyocamycin tetramer, each monomer as toyocamycin bound. **C.** AfRio1-adenosine tetramer, each monomer has adenosine bound.

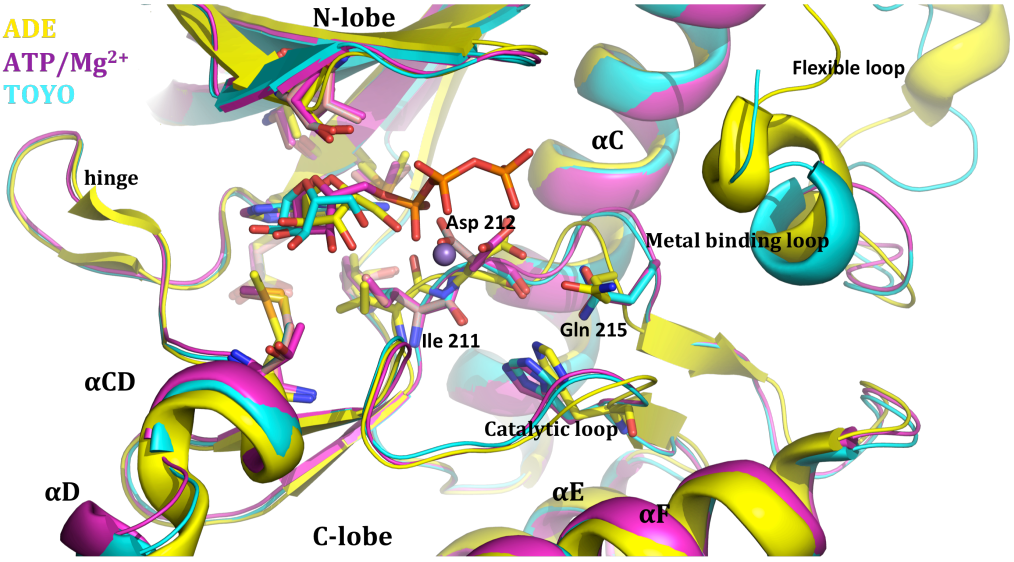


Figure 3.11: Alignment of afRio1 active site. Structural conformation switch observed in afRio1 bound to ATP (magenta), toyocamycin (cyan) and adenosine (yellow). The peptidyl carbonyl group of Ile 211 points towards the binding site in the adenosine complex whereas in toyocamycin and ATP this group points away from the binding site. This positions Asp 212 within hydrogen bonding distance to Gln 215 in the toyocamycin and ATP bound complexes. A shift is observed in the metal binding loop, the catalytic loop and the flexible loop. This conformation changes led to two different oligomers, toyocamycin and ATP stabilize the same type of oligomer whereas adenosine stabilizes a different type of oligomer.

.

Chapter 4: Structural Determination and Biochemical Studies of Human Rio1 Kinase (RioK1).

4.1 Introduction

Studies show that human Rio1 Kinase (RioK1) is over-expressed in colon cancer cells and recent reports link RioK1 to rRNA processing via recruitment of nucleolin to the human protein arginine methyltransferase complex (PRMT5) (*47, 53*). Nucleolin is a eukaryotic protein involved in ribosome biogenesis that is methylated by PRMT5 complex, thereby modulating nucleolin's interaction with rRNA (*47*). Immunoprecipitation and pull-down assays to identify potential interacting protein partners for RioK1 identified nucleolin and PRMT5. RioK1 interaction with PRMT5 is well documented as being stoichiometric and the interaction occurs via the N-terminal of PRMT5 and residues 1-242 of RioK1. (*47*) (*87*). Despite these discoveries, RioK1 substrates have not yet been identified, its biological function and mechanistic details are not fully understood and there are no reports of a eukaryotic Rio1 crystal structure. The availability of a RioK1 crystal structure would not only provide the first eukaryotic Rio1 structure but would also aid in addressing some of the unknowns.

RioK1 crystallographic studies were initially carried out with a RioK1 construct comprised of residues 106-494 that had been cloned by my advisor Dr. LeBlanc. Results from this construct initiated optimization studies that included limited proteolysis in order to determine a stable RioK1 construct and surface entropy

mutagenesis (SER) in order to obtain well-ordered, better diffracting crystals of RioK1 (106-494). SER targets surface residues found in clusters that have large hydrophilic side chains and replaces them with small amino acids such as alanines. Mutation of surface residues has the promise of generating x-ray quality crystals with some reports indicating an increase of resolution by more than 1 Å (*88, 89*). The rationale is that these surface modifications lead to reduction of conformational entropy and generate homogenous surface patches (*88*).

The free energy (ΔG) that drives crystallization is a sum of enthalpy of intermolecular interactions in the crystal (ΔH), protein entropic changes ($\Delta S_{protein}$) and solvent entropic changes ($\Delta S_{solvent}$) as shown in the equation below:

$$\Delta G = \Delta H - T (\Delta S_{protein} + \Delta S_{solvent})$$

Equation 4.1: Free energy that drives crystallization

Crystallization is more sensitive to both protein and solvent entropic changes and less so to enthalpic effects (*88*). Proteins with surfaces populated with large hydrophilic residues have side chains with high conformational entropy due to flexibility of the side chains. This inhibits crystallization by increasing the entropic barrier that must be overcome to order and bury these residues at points of crystal contacts (*90*) . In addition, these large hydrophilic residues have a large surface exposed to solvent preventing intermolecular contacts necessary for crystallization. Therefore replacing the large residues with smaller residues that have low conformational entropy favors the formation of crystal contacts and more ordered crystals (*90*). Mutational studies

have shown that lysine and glutamate are the best large hydrophilic targets whereas alanine is the most suitable small residue to replace Lys or Glu (*91*). The experimental and results of SER mutants for RioK1 will be described in more detail.

This chapter presents the first eukaryotic Rio1 crystal structure and compares the structure to the archaeal Rio1's crystal structure to determine whether evolution led to distinct structural features. In addition, biochemical studies including identification of autophosphorylation sites, substrate inhibition, affinity for potential Rio1 inhibitors and formation of oligomeric states are presented. Archaeal Rio1 (afRio1) studies revealed a high affinity for toyocamycin, substrate inhibition by high concentrations of ATP and formation of oligomeric states in the presence of different ligands. Similar biochemical studies were carried out using RioK1 to determine if any of these trends are in eukaryotic Rio1.

4.2 Experimental Procedures:

4.2a Expression and purification of RioK1 (106-494).

Initial attempts to express full-length RioK1 were not fruitful despite efforts to optimize the expression protocol by varying the concentration of Isopropyl β-D-1-thiogalactopyranoside (IPTG), trying expression at different temperatures and using different *E.coli* strains. A new construct of RioK1 residues 106-494 was cloned based on secondary structure prediction and incorporation of the RIO domain, residues 180 to 479. RioK1 (106-494) expression vector with a T7-promoter, an N-terminal histidine tag, a Tobacco Etch Virus (TEV) site and gene sequence was transformed to Rosetta™(DE3) pLysS *E.coli* cells (Novagen). The cells were plated on Lysogeny

Broth (LB) Agar plates containing 100 μg/mL ampicillin and 34 μg/mL chloramphenicol and incubated at 37°C for 16 hours. Colonies were inoculated into 40 mL of LB media with 100 μg/mL of carbenicilin and 34 μg/mL of chloramphenicol and grown overnight at 37°C while shaking at 250 rpm. The overnight culture was added to 4L of LB media with 100 μg/mL ampicillin and 34 μg/mL chloramphenicol and grown at 37°C while shaking at 250 rpm until the OD_{600} was between 0.6 and 0.8, the mid-log phase. These cells were induced with 1 mM IPTG at 20°C for 18 hours and then centrifuged at 6,000 rpm and 4°C for 30 minutes. The bacterial pellet was stored at -80°C and later resuspended in lysis buffer containing 50 mM Tris pH 8.0, 200 mM NaCl, 2 mM $MgCl_2$, 1 mM ATP, 0.2 % β-mercaptoethanol, 0.1 mg/mL DNAse and 0.3X Bugbuster . The resuspended cell pellet was gently stirred on ice for 45 minutes, transferred to pre-cooled centrifuge tubes and immediately centrifuged at 18,000 rpm and 4°C for 30 minutes. The supernatant was filtered and loaded onto a pre-equilibrated 5 mL HisTrap column (GE healthcare) attached to a Fast Protein Liquid Chromatography (FPLC) system at a flow-rate of 2.5 mL/min. The HisTrap column was washed with 10 column volumes of lysis buffer to remove any unbound protein. Bound RioK1 (106-494) protein was subjected to an imidazole gradient of 100 mM to 1 M and the protein eluted between 150 mM and 300 mM. Partially purified fractions were pooled together and transferred to a dialysis membrane to which 1 mg/mL Tobacco etch virus (TEV) protease, needed to cleave the histidine tag, was added. The protein was dialyzed overnight at 4°C into 50 mM Tris pH 8.0, 200 mM NaCl, 50 mM Tris pH 8.0, 1 mM ATP, 2 mM $MgCl_2$ and 0.2 % β-mercaptoethanol and then loaded onto a 5 mL

HisTrap column pre-equilibrated with lysis buffer. The HisTrap column was washed with lysis buffer to elute cleaved RioK1 (106-494) that was further purified by size exclusion chromatography using Superdex 200 (GE healthcare) pre-equilibrated in 50 mM Tris pH 8.0, 200 mM NaCl, 1 mM ATP, 2 mM $MgCl_2$, 10 % glycerol and 0.2 % β-mercaptoethanol. All purification steps were monitored by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) for purity (Figure 4.1).

4.2b Crystallization, data collection, processing and refinement of RioK1 (106-494)

Purified RioK1 (106-494) was concentrated to 25 mg/mL and subjected to robotic sparse matrix screening using Phoenix Liquid Handing System (Art Robbins). Commercially available matrix screens, Index (Hampton), Cyros (Qiagen), PEGs suite (Qiagen), Wizard I, II, III (Emerald Biosystems) were used for initial crystal screening in 96-well plates by the sitting drop method in three different protein to well solution ratios (1:2, 1:1, 2:1). Needle-shaped crystals were observed in PEGs Suite in a 2:1 protein to well solution ratio (Figure 4.2 A-B). These crystals grew in 0.2M $Ca(OAc)_2$, 20% PEG 400, pH 8.0 and optimized by varying both the concentration of precipitant, PEG 400, and pH to generate diffraction quality crystals. Optimized crystals were flash frozen in their well conditions with 20% (v/v) glycerol as a cryoprotectant. Data collection was done at 100K at the NE-CAT beamline at Argonne, IL. The best diffraction was to 3.50 Å and this data was indexed, integrated and scaled using HKL2000 (*71*). The crystals belonged to $P3_221$ space group and the unit cell dimensions were a=78.99, b=78.99, c=110.45 and α=90°, β=90° γ=120°. An

output file that is necessary to determine the quality of data collected based on redundancy, data completeness and R_{symm} values is shown in Table 4.1. Attempts to solve the structure by molecular replacement using afRio1 (PDB ID: 1ZP9) as a search model were not successful. This is in part due to the low sequence identity, 21%, shared between RioK1 and afRio1 as well as the low-resolution data. The alternative was to generate selenium–derivatized crystals (Se-Met) and solve the structure by single wavelength anomalous dispersion (SAD).

4.2c Expression, purification and data collection of Se-Met RioK1 (106-494)

A glycerol stock of RioK1 (106-494) was inoculated in 40 mL of LB media with 100 μg/mL of carbenicilin and 34 μg/mL of chloramphenicol and grown overnight at 37°C while shaking at 250 rpm. The overnight culture was centrifuged at 5,000 rpm for 20 minutes at 4°C. The bacterial pellet was resuspended with M9 selenomethionine growth media (Medicilon) that was prepared according to the manufacturer's protocol and then added to 4L of M9 SeMet media with 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. The cells were grown at 37°C while shaking at 250 rpm until the OD_{600} was close to 1.2. Inhibitory amino acids and selenomethionine were added to the cells according to the manufacturer's protocol and the cells were transferred to 20°C for 15 minutes. Cells were then induced with 1 mM IPTG at 20°C and grown for 18 hours. The following day the cells were centrifuged at 6,000 rpm and 4°C for 30 minutes. Purification of Se-Met RioK1 (106-494), crystal screening and optimization followed the same protocol as the native protein. Se-Met RioK1 (106-494) crystals were observed in the same conditions as

the native crystals.

Se-Met RioK1 (106-494) crystals had the same crystal form and space group as native crystals. The crystals diffracted to 3.50 Å and data was collected at one wavelength, 0.97915 Å (Figure 4.2 C). The cell dimensions were a=79.52, b=79.52, c=110.86 and α=90°, β=90° γ=120° with an overall redundancy of 2.0, an R_{symm} of 28% and 98% data completion (Table 4.2).

4.2d Surface Entropy Mutagenesis (SER)

The surface entropy mutagenesis prediction (SERp) server from UCLA was used to identify Lys 269 and Glu 270 as good targets for mutation in RioK1 (106-494) (*92*) . Primers for the double mutant RioK1 K269AE270A were designed and site-directed mutagenesis carried out using QuikChange lightning kit (Agilent) according to the manufacturer's instructions. Expression, purification and initial crystal screening of RioK1 K269AE270A followed the same protocol as that of RioK1 (106-494). Tiny needle-shaped crystals were observed in Wizard III screen and grew in 24% PEG 3350, 0.1 M HEPES pH 5.3 and 0.2 M $CaCl_2$ (Figure 4.3). These crystals were not reproducible and therefore could not be optimized. Crystals obtained from the initial screen did not diffract.

4.2e Limited Proteolysis

In order to determine a stable fragment of RioK1, the protein was subjected to proteolysis in of three different proteases; trypsin, thermolysin and elastase. For the trypsin and thermolysin reactions, 5 mg/mL of RioK1(106-494) in 100 mM Tris pH

8.0, 100 mM NaCl and 5 mM $CaCl_2$ was treated with 1 mg/mL of the protease at 37°C. For the elastase reaction 5 mg/mL of RioK1 in 200 mM Tris pH 8.0 was treated with 1 mg/mL of elastase at 37°C. All reactions were subjected to a time-based assay and stopped at 0, 10, 30, 60 minutes and 24 hours by denaturing the protease. In order to visualize the most stable fragment, all reactions were run on a preparative SDS-PAGE gel and the most stable fragment (25kDa) appeared after 10 minutes and was present after 24 hours (Figure 4.4). The band representing this fragment was cut from the gel and analyzed by in-gel trypsin digest followed by peptide identification using tandem mass spectrometry (tandem-MS/MS). Mass spectrometry analysis was done at the University of Maryland proteomics center by Dr. Yan Wang. The fragment was identified as residues 252-468, which does not incorporate the full RIO domain, residues 180-479 (Figure 4.5). From RioK1 (106-494) preparative gels, it was evident that this construct was subject to degradation; a band just below RioK1 (106-494) was consistently present in all purification steps (Figure 4.1). The next approach was to identify this fragment and clone the construct.

4.2f RioK1 (143-494): Cloning, expression, purification and crystallization.

The band representing RioK1 (106-494) degradation was cut from the gel and sent to Proteomics center for in-gel trypsin digest and analysis by tandem-MS/MS. Peptide identification data revealed that degradation was from the N-terminus of RioK1 (106-494) and the fragment comprised of residues 143 to 494 (Figure 4.6). RioK1 (143-494) was cloned using Gateway technologies (Invitrogen). Two gene-specific primers were designed for cloning; the sense primer 5'GAG AAC CTG TAC

TTC CAG GGC GAA GCG GAT ATG TAT CGC ATT AAA GCG -3' that contains a sequence for Tobacco Etch Virus (TEV) cleavage site in red, and the antisense primer 5' **GGGG ACC ACT TTG TAC AAG AAA GCT GGG TT**A TCA GGT GCG TTC TTC CAC CTG GTT TTC 3' that contains a recombinant site, attb2, shown in bold. These primers along with the full-length RioK1 DNA were subjected to a polymerase chain reaction (PCR) after which the PCR product was subjected to a second reaction using a "piggyback primer" (Invitrogen) 5'-**GGGG ACA AGT TTG TAC AAA AAA GCA GGC TCG** GAG AAC CTG TAC TTC CAG that contains the second recombinant site, attb1, in bold as well as a TEV site in red. The final PCR product was introduced into an entry vector, pDONR 201 (Invitrogen), and finally into a destination vector, Pdest 527 (Invitrogen) that contains a T7 promoter and an N-terminal 6X histidine tag sequence. RioK1 (143-494) expression vector was transformed to Rosetta (DE3) pLysS *E.Coli* cells (Novagen) and expression carried out in SelenoMethionine expression media (Shanghai Medicilon Inc) according to the manufacturer's instructions. Purification of Se-Met RioK1 (143-494) followed the same protocol as RioK1 (106-494). Purified Se-Met RioK1 (143-494) with ATP/Mg^{2+} bound was concentrated to 20 mg/mL and subjected to robotic sparse matrix screening using Phoenix Liquid Handling system (Art Robbins) (Figure 4.7). Rod-shaped crystals were observed in Wizard III screen in both the 1:1 and 2:1 protein to well solution ratios. These crystals grew in 28% PEG 400, 0.1M HEPES pH 7.5, 0.2M $CaCl_2$ and were optimized by varying the concentration of precipitant, PEG 400, and the pH (Figure 4.8). Optimized crystals were flash frozen in their well solution with 20% (v/v) ethylene glycol as a cryoprotectant and shipped to Argonne

National Labs for data collection and processing.

4.2g Data collection, processing and refinement of RioK1 143-494

Data was collected at one wavelength (0.97915 Å), indexed, integrated and scaled using the HKL2000 package program as with previous crystals (*71*). The best diffraction was to 2.63 Å and the space group was P32. The unit cell dimensions were a=78.81Å, b=78.81Å, c= 110.57Å, α=90.00^0, β=90.00^0, γ=120.00^0 an overall redundancy of 4.5, 95.9% data completion and an overall R_{symm} of 10.6% (Table 4.5). The structure was solved using an automated experimental phasing program, AutoSol, that's part of PHENIX (*93*). AutoSol finds heavy atom sites using hybrid substructure search (HYSS) then calculates experimental phases using Phaser and improves the phases by density modification using RESOLVE (*94-96*). After obtaining density-modified phases the program carries out a round of model building and refinement. The initial RioK1 (143-494) model was subjected to two rounds of model building using AUTOBUILD that's part of PHENIX (*97-100*). Subsequent rounds of model building were carried out in COOT accompanied by rounds of refinement using PHENIX and REFMAC5 that's part of CCP4i (*74-76*). Waters were added to the structure using COOT and the final model had a R_{work}/R_{free} of 22.8/26.6 (Table 4.5). The refined crystal structure had two molecules per asymmetric unit (ASU) (Figure 4.9).

4.2h RioK1 autophosphorylation sites

A sample of purified RioK1 (106-494) was first analyzed for mass determination by electrospray ionization (ESI) and the results revealed three peaks on the deconvoluted spectrum (Figure 4.14). The first peak corresponded to the expected molecular weight of 45126.18 Daltons, the second and third peaks corresponded to singly (45203.69 Daltons) and doubly (45279.67 Daltons) phosphorylated RioK1 (106-494) respectively (Figure 4.14). In order to identify the autophosphorylation sites, the RioK1 sample was subjected to trypsin digest followed by tandem-MS/MS for peptide identification. This work was done in collaboration with Dr. Yan Wang at the University of Maryland Proteomics center. For analysis the expected molecular weight (MW) of various peptide fragments was compared to the observed MW in order to determine fragments that contained phosphorylation sites. One site, serine 130 was identified with certainty but it was ambiguous whether the second site was serine 415 or serine 416 since both were on the same peptide fragment. To address this ambiguity, Ser 415 and Ser 416 were mutated to alanines followed by autophosphorylation assays to determine the second site.

4.2i Autophosphorylation sites: Mutagenesis and Assay

Primers for mutants RioK1 (106-494) S130A, S415A and S416A were designed and site-directed mutagenesis was carried out using QuikChange™ Lightning kit (Agilent) according to manufacturer's instructions (Table 4.6). Three single mutants S130A, S415A and S416A and two double mutants S130AS415A and S130AS416A were expressed and purified following the protocol for wild-type

RioK1 (106-494) with one exception, 1 mM ATP was substituted for 500 μM adenosine in all the buffers. Adenosine was used in these purifications to prevent any autophosphorylation activity before the actual assays. Autophosphorylation assays were carried out by adding 10 μg of wild-type RioK1 (106-494) and mutants to a 30 μl reaction volume in a buffer containing 50 mM Tris pH 8.0, 200 mM NaCl, 500 μM adenosine, 2 mM $MgCl_2$ and 0.2 % β-mercaptoethanol. 1 μM of γ-^{32}P labeled ATP was added to all reactions and incubated at 25°C for 1 hour. The reactions were done in triplicates and stopped by adding 8 μL of SDS loading buffer and heating at 95°C for 5 minutes to denature the protein. All reactions were run on a NuPAGE™ 4-12% Bis-Tris denaturing gel (Invitrogen) to separate the protein from γ-^{32}P labeled ATP. The gel was dried, exposed onto a phosphorimager screen for 15 hours and quantified using ImageQuant to determine autophosphorylation activity based on amount of **γ-**^{32}phosphate incorporated into the protein (Figure 4.15 A).

4.2j RioK1 Thermofluor Assay

A Thermofluor assay was used to test the affinity of RioK1 for toyocamycin, sangivamycin, 7-methylinosine and ATP/Mg^{2+}. 1.0 mg/mL of purified RioK1 was incubated at 4°C for 1 hour with 2 mM of toyocamycin, sangivamycin, 7-methylinosine and ATP/Mg^{2+} in a buffer containing 50 mM Tris pH 8.0, 200 mM NaCl, 2 mM $MgCl_2$, 0.2 % β-mercaptoethanol and 1 mM ATP. The control reaction was RioK1 with no ligand added. After incubation, Sypro-Orange™ fluorescent dye (Sigma-Aldrich) was added to the reactions and heated from 4°C to 98°C at a rate of 0.2°C per second using Bio-Rad Mini-Opticon™ thermal cycler. Measurements were

done in triplicates and the T_m was determined as the inflection point of the melting curve. The change in T_m was calculated as the difference between the T_m of bound RioK1 and T_m of unbound RioK1.

4.2 k Substrate inhibition at high concentrations of ATP

An autophosphorylation assay was carried out to test whether RioK1 activity is inhibited by high concentrations of ATP. In this assay, 5 μg of RioK1 was added to a buffer containing 20 mM Tris pH 8.0, 100 mM NaCl, and 2 mM $MgCl_2$. The concentration of γ-^{32}P labeled ATP was varied from 1800 μM to 0.9 μM. All RioK1 reactions were incubated at 25°C for 1 hour and components separated by running a denaturing gel. Autophosphorylation activity was determined by quantifying the gels using ImageQuant.

4.2l Size Exclusion chromatography

Size exclusion chromatography (SEC) was used to determine oligomeric states in the presence of 2 mM adenosine and ATP/Mg^{2+} as well as phosphorylated RioK1. All samples had a final protein concentration of 3 μg/μl and were analyzed using an analytical Superdex™ 200 column (GE Healthcare) that was pre-equilibrated with a buffer containing 50 mM Tris pH 8.0, 200 mM NaCl, 2 mM $MgCl_2$ and 0.2 % β-mercaptoethanol. In order to calculate the estimated molecular weight of RioK1 samples, calibration proteins (GE healthcare); ribonuclease (13kDa), ovalbumin (43kDa), conalbumin (75kDa) and aldolase (158kDa) were also analyzed by SEC. Calibration proteins were diluted with 50 mM Tris pH 8.0 and 200 mM NaCl for a

final concentration of 3 μg/μl, mixed together, centrifuged to remove any precipitants and loaded onto the analytical Superdex™ 200 column.

4.3 Results

4.3a Se-Met RioK1 (106-494)

Attempts to solve RioK1 (106-494) by SAD phasing were not successful due to a low anomalous signal as a result of radiation damage during data collection and possible oxidation of selenium atoms. In addition the data set should have been cut back to about 3.9 Å during scaling since beyond this resolution, the R_{symm} is above 0.5, which is generally the cut-off fraction and inclusion of this data resulted in the high overall R_{symm} value of 28%. The redundancy was also below the acceptable value of 3.0 and above. This is usually resolved by collecting more data as long as the crystal does not succumb to radiation damage. The resulting electron density map lacked enough detail to properly model in the protein. After several attempts at model building and refinement in PHENIX and COOT, the R_{work}/R_{free} did not dropped below 41.1/45.5% indicative of poor agreement between the model and the original x-ray data (*75, 76*). Se-Met RioK1 (143-494) with ATP/Mg^{2+} resulted in better diffracting crystals that were used to solve the first crystal structure of a eukaryotic Rio1 kinase.

4.3 b Overall Structure of RioK1 (143-494) with ATP/Mg^{2+} bound.

The overall RioK1 (143-494)-ATP /Mg^{2+} structure consists of an N-lobe that has the helix αR, and a twisted β sheet that contains the nucleotide binding loop (p-loop) whereas the C-lobe is helical and houses the catalytic and metal binding loops

(Figure 4.9). Both N- and C-lobes are connected by a hinge region that buries the adenine ring of ATP. The RioK1 (143-494)-ATP /Mg^{2+} overall structure is consistent with the general features of kinases that contain an N-lobe comprised of a twisted β-sheet and a helical C-lobe (*10, 29*).

Electron density is not observed for residues 143-163 that corresponds to the subfamily-specific N-terminal helices (α1 and α2) and on the C-lobe there is no density for residues 431-494 on chain A and residues 429-494 on chain B. Lack of electron density on the N- and C- terminus is due to disorder and flexibility of these regions. The structure also lacks electron density for residues 212-243 that consists of the flexible loop. This is similar to the structure of afRio1-ATP/Mg^{2+} complex for which no electron density is observed for residues 85-91 that's part of the flexible loop (*24*). An alignment of the two molecules of RioK1 (143-494)-ATP /Mg^{2} in the asymmetric unit revealed an RMSD of 0.237 Å over 192 residues (Figure 4.10).

An alignment of afRio1-ATP/Mn^{2+} and RioK1 (143-494)-ATP/Mg^{2+} dimer structures had an RMSD of 0.894 Å over 158 residues and shifts on the N-lobe and the metal binding loop on the C-lobe were observed (Figure 4.11A). The N-lobe helix, αR, on RioK1 (143-494) is shifted 3.0Å away from afRio1's αR resulting in a shift on the N-lobe twisted β sheet and the nucleotide-binding loop (p-loop) (Figure 4.11). Subsequently shifts are observed on the hinge region and the metal binding loops (Figure 4.11). RioK1 (143-494) metal binding loop is shifted 1.1 Å away from afRio1's metal binding loop. In addition, the RioK1 (143-494)-ATP complex shows the γ-phosphate of ATP positioned and orientated differently compared to what's observed in the afRio1-ATP complex.

4.3 c RioK1 (143-494) interactions with ATP

The RioK1 (143-494)-ATP /Mg^{2+} structure shows ATP buried in a pocket that features hydrophobic contact with Ile 186, Val 194, and Ala 206 from the N-lobe; Met 277, Phe 279, Ile 280 and Pro 286 from the hinge region; Leu 331 and Ile 340 on the metal-binding loop, from the C-lobe (Figure 4.12A). These hydrophobic interactions are very similar to the ones observed in the afRio1-ATP/Mn^{2+} complex (*24*). In addition to hydrophobic contacts, ATP participates in several hydrogen bond interactions that are either direct or water-mediated. There is a direct hydrogen bond from the peptidyl carbonyl oxygen of Ser 278 on the hinge region to the amine group at N6 position on the adenine ring (Figure 4.12B). In the afRio1-ATP/Mn^{2+} structure this interaction is between the amine group at N6 and the peptidyl carbonyl oxygen of Glu 148 (*24*). Two hydrogen bonds are observed; one from the indole nitrogen (N1) to the peptidyl amine group of Ile 280 on the hinge (Figure 4.12B) and another bond between Lys 208 and an α-phosphate oxygen. Similar interactions are observed in the afRio1-ATP complex (*24*).There is a metal-mediated interaction between the catalytic residue, Asn 329, to the α, β and γ-phosphate oxygens of ATP (Figure 4.12B). There is an additional metal-mediated interaction between Mg^{2+} to the metal binding residue, Asp 341, to the α, β and γ-phosphate oxygens of ATP (Figure 4.12B). Both metal-mediated interactions are similar to the ones observed in the afRio1-ATP/Mn^{2+} complex with the metal binding residue Asp 212 and the catalytic residue Asn 201 (*24*). Water-mediated hydrogen bonds are observed between Ser 187 of the p-loop and a β-phosphate oxygen of ATP and between Glu 191 of the p-loop and the β- and

γ-phosphate oxygens of ATP (Figure 4.12B). A key difference between RioK1 (143-494)-ATP/Mg^{2+} and afRio1-ATP/Mn^{2+} complexes is the distance between the catalytic aspartate and the γ-phosphate oxygen of ATP. In the RioK1 (143-494)-ATP /Mg^{2+} this distance is 3.3 Å whereas in afRio1-ATP/Mn^{2+} this distance is 5.1 Å. The difference in this distance can be attributed to the shift in the metal binding loop as well as the orientation and positioning of the γ-phosphate. This is suggestive of a difference in the catalytic rate activity between eukaryotic Rio1 and afRio1. Therefore despite evolution, both eukaryotic RioK1 and archaeal Rio1 share almost identical interactions with ATP.

4.3d Potential candidates for dimer interfaces

The computational database, PDBe PISA (protein interfaces, surfaces and assemblies) server, was used to analyze RioK1 (143-494)-ATP/Mg^{2+} potential interacting interfaces and calculate the buried surface area. The largest interface had a buried surface area of 564.3 $Å^2$ per molecule and is observed between the two molecules in the asymmetric unit. This interface is characterized by interactions between the hinge region of monomer A and the C-lobe residues of monomer B (Figure 4.13A). This interface is characterized by poor packing and only two hydrogen bonds are observed, one between Lys 401 of monomer A and Asp 284 of monomer B and another between Asp 387 of monomer A and Lys 282 of monomer B (Figure 4.13A). The second potential interface is between symmetry related molecules and has a calculated surface area of 429.1 $Å^2$ per molecule. The interface is observed between αC on the N-lobe of molecule A and C-lobe residues of molecule

B (Figure 4.13B). This interface is characterized by hydrophobicity, a characteristic of biological interface, and better packing than the first interface between two molecules per asymmetric unit. This second interface has hydrogen interactions between: Glu 413 of A and Arg 319 of B; His 419 of A and Thr 259 of B; Gln 417 of A and Arg 252 of B; and Arg 427 of A and peptidyl carbonyl oxygens of Arg 271 of molecule B (Figure 4.13 B). Hydrophobic contacts are observed between Val 421 of A and Ile 255 of B; Val 424 of A and Met 251 of B; and Ala 428 of A and Leu 270 of B. Both interfaces are different from the afRio1-ATP dimer interface that positions the helices of the C-lobe of one monomer in contact with both the N- and C-lobes of the second monomer almost completely occluding the active site.

4.3e Identification of RioK1 autophosphorylation sites.

Rio kinases that have been successfully expressed and purified thus far have serines as sites of autophosphorylation. One of the goals was to identify these autophosphorylation sites in RioK1 since there are no such reports to date. From the radioactivity gels the mutant S130A had diminished autophosphorylation activity as expected and the double mutant S130AS416A lacked any activity (Figure 4.15 A-B). This confirmed that serine 416 was the second autophosphorylation site. Both Ser 130 and Ser 416 are not found on the flexible loop, as is the case with afRio1 that has its autophosphorylation site, Ser 108, on the flexible loop.

4.3f RioK1 Affinity for Toyocamycin.

Unbound RioK1 had a T_m of 41.1 ±1.1°C whereas RioK1 with toyocamycin bound had a T_m of 53.2 ±1.8°C corresponding to a 12.1°C shift (Figure 4.16). A similar T_m shift of 12.0°C is observed with afRio1 bound to toyocamycin. RioK1 with either sangivamycin or 7-methylinosine bound, had a T_m of 50.7°C, a 9.6°C increase and in the presence of ATP/Mg^{2+} the T_m was 50.0 ±1.2°C, a 8.9°C shift (Figure 4.16). These results confirmed that toyocamycin binds to RioK1 with a higher affinity than ATP/Mg^{2+}, since it stabilizes RioK1 more than ATP/Mg^{2+}. The same trend was observed with afRio1 supporting conservation between RioK1 and afRio1 in regards to affinity for toyocamycin. This similarity between afRio1 and RioK1 is somewhat expected considering that both proteins interact with ATP in an almost identical fashion and therefore RioK1 might have similar interactions with toyocamycin as those observed between afRio1 and toyocamycin. Obviously a crystal structure of RioK1 with toyocamycin bound would be more telling.

4.3g Substrate inhibition by high concentrations of ATP.

At high concentrations of ATP, autophosphoryation activity is diminished and is almost unmeasurable at 1800 μM and activity is only detectable at lower ATP concentrations as observed on the radioactive gels (Figure 4.17). This is consistent with afRio1 studies that showed inhibition of afRio1 activity at high concentrations of ATP, suggesting that the two kinases may share a similar substrate inhibition mechanism that involves formation of oligomeric states that are promoted by high

concentrations of ATP. In this mechanism, inhibition occurs via modulation of binding sites in oligomers that are triggered by high concentrations of ATP.

4.3h: Formation of Oligomers

Initial attempts to analyze oligomerization by sedimentation equilibrium resulted in aggregation of RioK1 and therefore size exclusion chromatography was used for analysis. Calibration proteins were used to generate a plot of Log of molecular weight as a function of elution volume was fit to linear regression using GraphPad prism. This plot was used to calculate the estimated molecular weight of the different RioK1 samples. The monomeric molecular weight of RioK1 is 45.13 kDa, the phosphorylated RioK1 had a calculated molecular weight of 70.3 kDa and eluted as a mixture of monomer and dimer (Figure 4.18). In the presence of ATP/Mg^{2+} RioK1 eluted mostly as a dimer with a calculated molecular weight of 104.3 kDa and in the presence of adenosine, RioK1 eluted as a tetramer with a calculated molecular weight of 179.1 kDa (Figure 4.18). Taken together this data shows that RioK1 does form higher order oligomers in the presence of ligands, consistent with observations made with afRio1. In addition, phosphorylation of RioK1 that occurs near the vicinity of two symmetry-related molecules, promotes some monomerization consistent with results from phosphorylated afRio1 (Figure 4.13B). This suggests that both proteins can access multiple oligomeric states and could employ the same type of substrate inhibition mechanism and that phosphorylation may regulate catalytic activity. Steady state kinetic experiments of RioK1 would validate this hypothesis.

4.4 Discussion

A substantial amount of work went towards structural determination of RioK1 from trying different constructs to applying limited proteolysis and surface entropy mutagenesis. Comparison between the crystal structure of RioK1-ATP complex and that of afRio1-ATP reveals that both complexes share similar interactions with ATP despite evolution. Differences between the two complexes include: a) the distance between the catalytic aspartate and the γ-phosphate of ATP, and b) crystal contacts/ interacting interfaces. Shifts are also observed on the N-lobe that houses the p-loop, the hinge region and metal binding loops of RioK1 relative to afRio1.

The fact that both RioK1 and afRio1 share similar high binding affinities to toyocamycin may be an indication that both proteins bind to toyocamycin in a similar fashion and may maintain similar interactions. A co-crystal structure of RioK1 and toyocamycin would aid in characterizing the binding. In addition, toyocamycin can be used as a tool to identify RioK1 substrates in eukaryotic systems using methods described in chapter 2. Inhibition of RioK1 by toyocamycin can be carried out using kinase reactions similar to the ones used in the aracheal system. Other inhibition studies would be to determine the effect of toyocamycin on cell-cycle progression since eukaryotic Rio1 is implicated in the progression of cells from the G1 phase to the S phase and in the metaphase.

This chapter presents the first eukaryotic structure of Rio1 that not only aided in identifying interactions of the protein with ATP but also potential biological

interfaces. It would be of interest to test how mutating residues involved in the interacting interfaces would affect dimerization that is observed in both crystallization and solution.

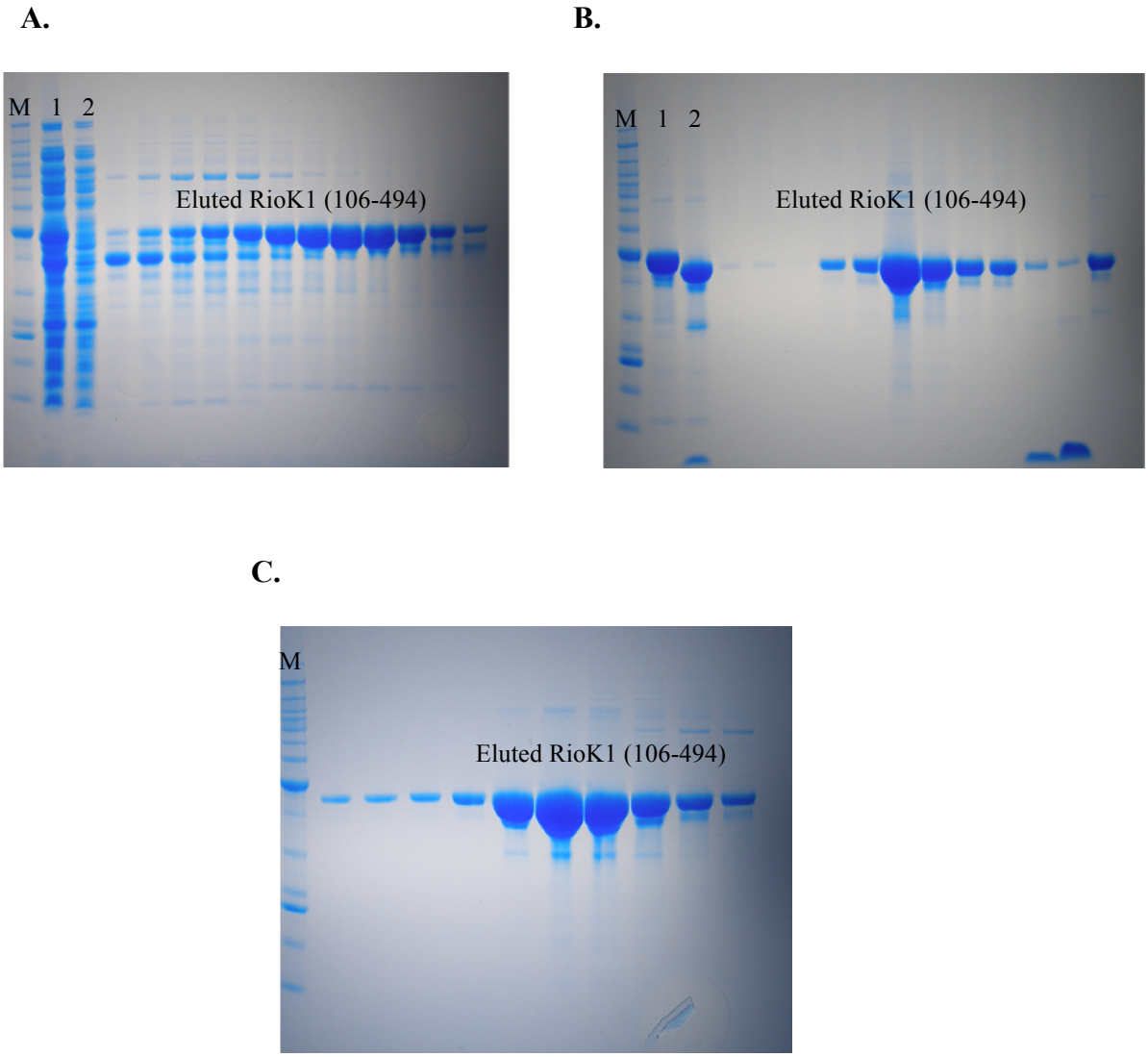


Figure 4.1: Preparative gels of RioK1 (106-494) purification. Expected molecular weight is 45.13 kDa. Lane M is the protein ladder **A.** Gel after first HisTrap column purification, lane 1 is the lysate and lane 2 is the flow-through after loading the protein onto the column. **B.** Gel after the second HisTrap column purification, lane 1 is before histidine tag cleavage and lane 2 is after cleavage. **C.** Gel after size exclusion chromatography.

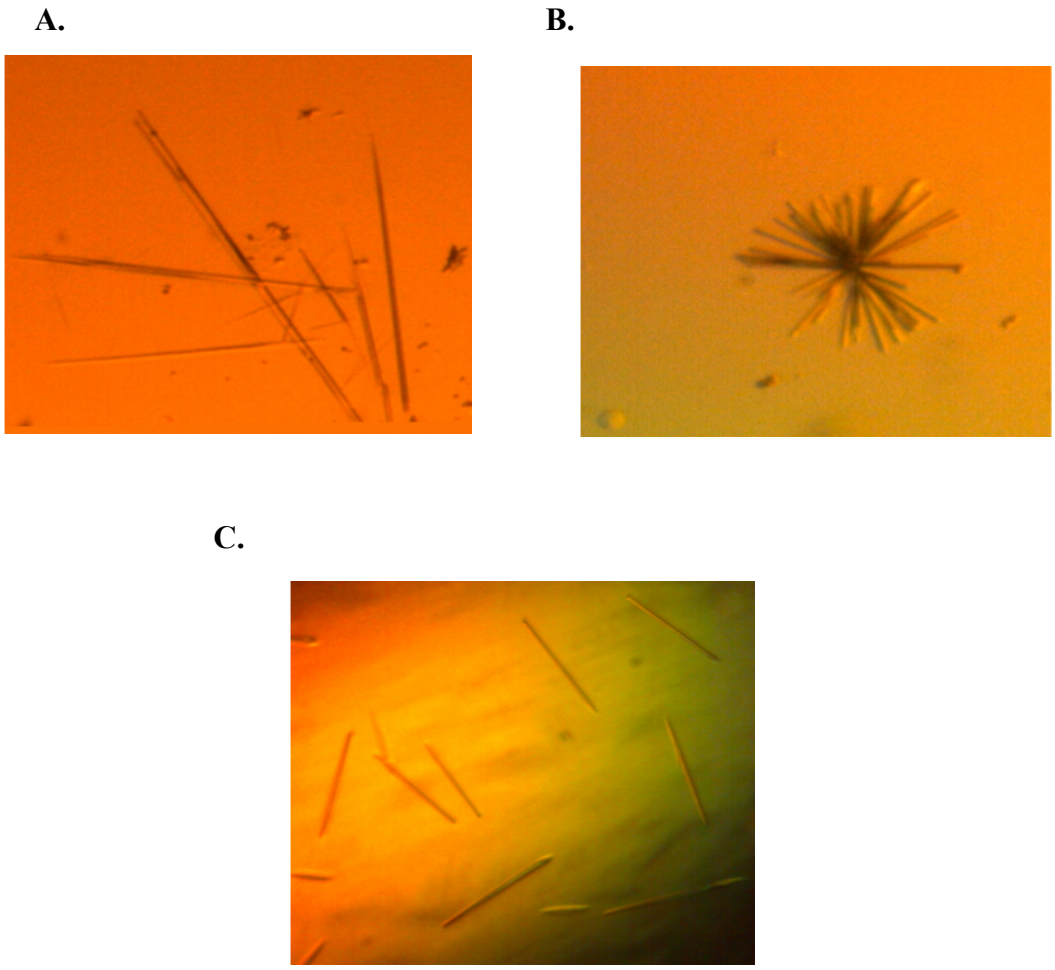


Figure 4.2: RioK1 (106-494) crystals. **A.** Needle-shaped crystals in PEGs Suite screen before optimization. **B.** RioK1 (106-494) optimized crystals that grew in 0.2 M Calcium acetate and 20% PEG 3350. **C.** Se-Met RioK1 (106-494) crystals in PEGs Suite screen.

Table 4.1: Scalepack log file: RioK1 (106-494)

Shell Lower limit	Shell Upper limit	Average Redundancy Per Shell
50.00	8.00	2.6
8.00	6.36	2.8
6.36	5.55	2.8
5.55	5.05	2.8
5.05	4.69	2.7
4.69	4.41	2.6
4.41	4.19	2.5
4.19	4.01	2.4
4.01	3.85	2.5
3.85	3.72	2.4
3.72	3.60	2.3
3.60	3.50	2.4
All hkl		2.6

Shell Lower limit	Shell Upper limit	I/Sigma in resolution shells: % of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
50.00	8.00	0.4	1.2	3.2	5.1	7.2	15.1	33.2	66.4	99.6
8.00	6.36	2.6	8.4	14.8	21.4	33.2	55.0	78.7	21.2	99.9
6.36	5.55	6.9	18.2	29.5	40.6	55.5	75.3	92.6	7.3	99.9
5.55	5.05	7.6	18.6	34.2	45.6	59.3	76.9	89.5	10.5	100.0
5.05	4.69	10.3	22.9	34.4	45.8	60.6	78.7	92.0	8.0	100.0
4.69	4.41	10.5	26.6	41.8	53.4	64.0	78.3	89.3	9.0	98.3
4.41	4.19	14.1	32.4	47.7	56.7	67.7	81.2	88.7	5.2	93.9
4.19	4.01	15.6	35.1	51.3	60.2	69.6	77.2	82.8	5.4	88.3
4.01	3.85	17.1	34.3	46.4	53.0	59.8	67.4	73.3	4.1	77.3
3.85	3.72	16.4	35.6	48.3	53.9	58.5	64.6	68.4	3.9	72.3
3.72	3.60	15.5	32.8	46.4	51.5	56.9	61.4	64.1	2.4	66.6
3.60	3.50	14.7	30.5	39.6	43.8	47.8	52.2	56.5	0.6	57.1
All hkl		11.0	24.7	36.4	44.2	53.3	65.3	75.7	12.0	87.8

Shell limit	Lower Angstrom	Upper	Average I	Average error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac
50.00	8.00	290.6	8.8	6.2	2.840	0.051	0.056	
8.00	6.36	76.8	5.0	4.7	1.871	0.089	0.078	
6.36	5.55	37.0	4.9	4.8	1.551	0.167	0.130	
5.55	5.05	41.8	5.8	5.7	1.682	0.178	0.133	
5.05	4.69	41.1	6.4	6.3	1.719	0.186	0.128	
4.69	4.41	49.4	7.3	7.1	1.849	0.164	0.117	
4.41	4.19	33.8	7.6	7.5	1.635	0.210	0.148	
4.19	4.01	34.4	8.4	8.3	1.738	0.201	0.131	
4.01	3.85	30.1	8.4	8.3	1.717	0.240	0.157	
3.85	3.72	30.2	8.9	8.8	2.494	0.253	0.157	
3.72	3.60	23.3	9.3	9.2	1.615	0.278	0.170	
3.60	3.50	16.5	9.3	9.2	1.821	0.443	0.359	
All reflections		63.4	7.3	6.9	1.869	0.134	0.088	

Table 4.2: Scalepack log file: Se-Met RioK1 (106-494)

Shell Lower limit	Shell Upper limit	Average Redundancy Per Shell
50.00	7.53	2.1
7.53	5.98	2.0
5.98	5.23	2.1
5.23	4.75	2.0
4.75	4.41	2.1
4.41	4.15	2.1
4.15	3.94	2.0
3.94	3.77	2.0
3.77	3.63	2.0
3.63	3.50	1.9
All hkl		2.0

Shell Lower limit	Shell Upper limit	I/Sigma in resolution shells: % of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
50.00	7.53	0.9	2.8	5.3	8.7	16.2	32.9	63.7	32.8	96.5
7.53	5.98	7.4	17.2	29.8	40.5	56.9	79.4	94.0	5.0	98.9
5.98	5.23	12.7	28.8	45.3	59.5	75.6	89.7	97.2	1.3	98.6
5.23	4.75	13.0	28.5	46.8	60.3	76.1	90.8	97.3	1.5	98.8
4.75	4.41	14.1	31.0	49.1	62.7	75.2	88.4	96.2	2.6	98.7
4.41	4.15	20.2	39.8	58.7	72.0	83.5	93.2	97.5	1.1	98.6
4.15	3.94	21.6	44.5	65.6	77.3	86.4	93.3	97.8	0.9	98.7
3.94	3.77	26.0	52.2	71.6	81.5	88.8	94.5	97.2	0.8	98.0
3.77	3.63	31.9	58.5	76.5	85.4	91.1	94.5	96.9	0.6	97.5
3.63	3.50	31.1	58.9	77.7	86.6	91.6	94.5	95.7	0.3	96.0
All hkl		17.9	36.2	52.6	63.4	74.1	85.1	93.3	4.7	98.0

Shell limit	Lower Angstrom	Upper Angstrom	Average I	Average error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac
50.00	7.53		61.9	2.9	2.5	5.125	0.084	0.098
7.53	5.98		15.3	2.1	2.0	2.687	0.178	0.130
5.98	5.23		8.7	2.3	2.2	2.102	0.316	0.233
5.23	4.75		10.4	2.7	2.7	2.206	0.324	0.171
4.75	4.41		13.2	3.1	3.1	2.716	0.328	0.206
4.41	4.15		8.5	3.4	3.3	2.327	0.490	0.328
4.15	3.94		8.7	3.7	3.7	2.443	0.567	0.289
3.94	3.77		7.2	4.1	4.0	2.839	0.707	0.701
3.77	3.63		5.9	4.3	4.2	2.220	0.807	0.351
3.63	3.50		4.1	4.5	4.4	1.842	0.950	0.384
All reflections			14.3	3.3	3.2	2.667	0.280	0.186

A. **B.**

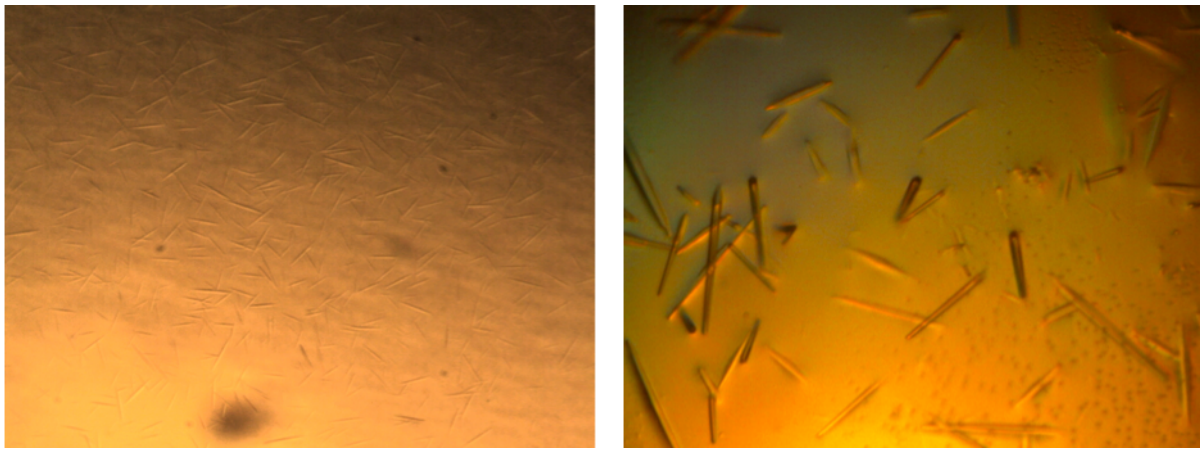


Figure 4.3: Surface Entropy Mutagenesis (SER), RioK1_K269AE270A crystals.

A. Crystals observed in Wizard III screen. **B.** Optimized crystals that grew in 24% PEG 3350, 0.1 M HEPES pH 5.3 and 0.2 M $CaCl_2$.

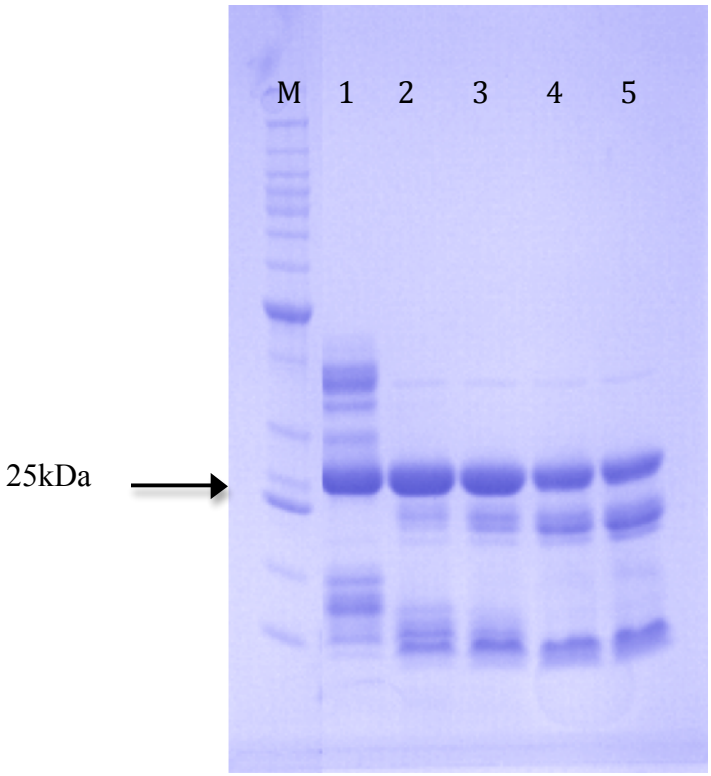


Figure 4.4: Limited Proteolysis assay. The gel represents trypsin proteolysis reaction. Lane M is the protein ladder. Lanes 1, 2, 3, 4 and 5 are samples after 0, 10, 30, 60 minutes and 24 hours. The band at 25kDa is observed after 10 minutes and is present after 24 hours.

Protein View

Match to: **RIOK1_HUMAN** Score: **3699**
Serine/threonine-protein kinase RIO1 OS=Homo sapiens GN=RIOK1 PE=1 SV=2
Found in search of 20100115ID02.mgf

Nominal mass (M_r): **65884**; Calculated pI value: **5.84**
NCBI BLAST search of RIOK1_HUMAN against nr
Unformatted sequence string for pasting into other applications

Taxonomy: Homo sapiens

Fixed modifications: Carbamidomethyl (C)
Variable modifications: Phospho (S),Phospho (T),Phospho (Y)
Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
Sequence Coverage: **42%**

Matched peptides shown in **Bold Red**

```
     1 MDYRRLLMSR VVPGQFDDAD SSDSENRDLK TVKEKDDILF EDLQDNVNEN
    51 GEGEIEDEEE EGYDDDDDDW DWDEGVGKLA KGYVWNGGSN PQANRQTSDS
   101 SSAKMSTPAD KVLRKFENKI NLDKLNVTDS VINKVTEKSR QKEADMYRIK
   151 DKADRATVEQ VLDPRTRMIL FKMLTRGIIT EINGCISTGK EANVYHASTA
   201 NGESRAIKIY KTSILVFKDR DKYVSGEFRF RHGYCKGNPR KMVKTWAEKE
   251 MRNLIRLNTA EIPCPEPIML RSHVLVMSFI GKDDMPAPLL KNVQLSESKA
   301 RELYLQVIQY MRRMYQDARL VHADLSEFNM LYHGGGVYII DVSQSVEHDH
   351 PHALEFLRKD CANVNDFFMR HSVAVMTVRE LFEFVTDPSI THENMDAYLS
   401 KAMEIASQRT KEERSSQDHV DEEVFKRAYI PRTLNEVKNY ERDMDIIMKL
   451 KEEDMAMNAQ QDNILYQTVT GLKKDLSGVQ KVPALLENQV EERTCSDSED
   501 IGSSECSDTD SEEQGDHARP KKHTTDPDID KKERKKMVKE AQREKRKNKI
   551 PKHVKKRKEK TAKTKKGK
```

Figure 4.5: Mass spectrometry analysis for 25kDa fragment. Peptide identification from limited proteolysis assay. Peptide fragments in red were identified and the 25kDa fragment includes residues 252-468.

Rio1 coverage for the second lower band. Amino acids identified: 143-493

1 MDYRRLLMSR VVPGQFDDAD SSDSENRDLK TVKEKDDILF EDLQDNVNEN
51 GEGEIEDEEE EGYDDDDDDW DWDEGVGKLA KGYVWNGGSN PQANRQTSDS
101 SSAKMSTPAD KVLRKFENKI NLDKLNVTDS VINKVTEKSR QK**EADMYRIK**
151 **DKADRATVEQ VLDPRTRMIL FKMLTRGIIT EINGCISTGK EANVYHASTA**
201 **NGESR**AIK**IY KTSILVFKDR DKYVSGEFR**F RHGYCKGNPR KMVKTWAEKE
251 MRNLIR**LNTA EIPCPEPIML RSHVLVMSFI GKDDMPAPLL KNVQLSESK**A
301 R**ELYLQVIQY MR**RMYQDAR**L VHADLSEFNM LYHGGGVYII DVSQSVEHDH**
351 **PHALEFLRKD CANVNDFFMR HSVAVMTVRE LFEFVTDPSI THENMDAYLS**
401 **K**AMEIASQRT KEER**SSQDHV DEEVFKR**AYI PR**TLNEVKNY ERDMDIIMKL**
451 **KEEDMAMNAQ QDNILYQTVT GLK**KDLSGVQ K**VPALLENQV EER**TCSDSED
501 IGSSECSDTD SEEQGDHARP KKHTTDPDID KKERKKMVKE AQREKRKNKI
551 PKHVKKRKEK TAKTKKGK

Figure 4.6: Mass Spectrometry analysis for RioK1 (143-494). Peptide fragment representing RioK1 (106-494) degradation is identified as residues 143-494, shown in red.

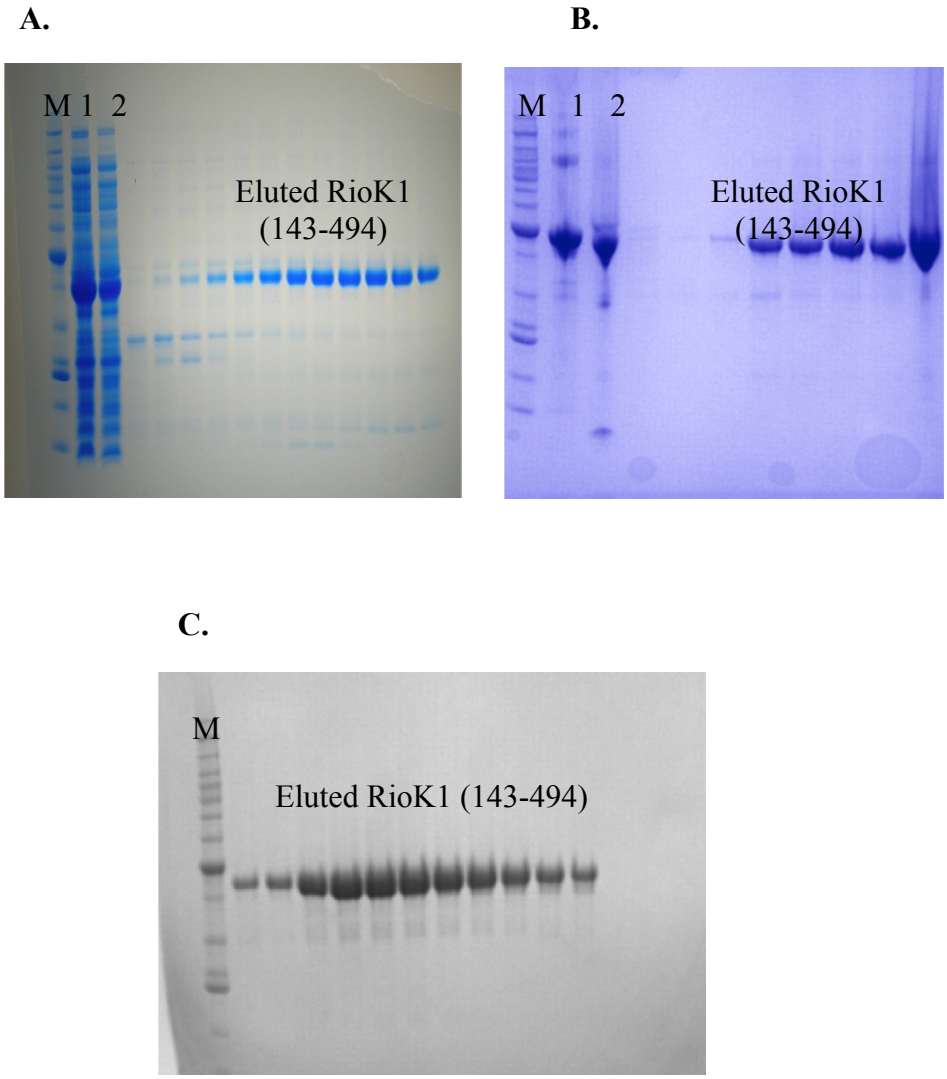
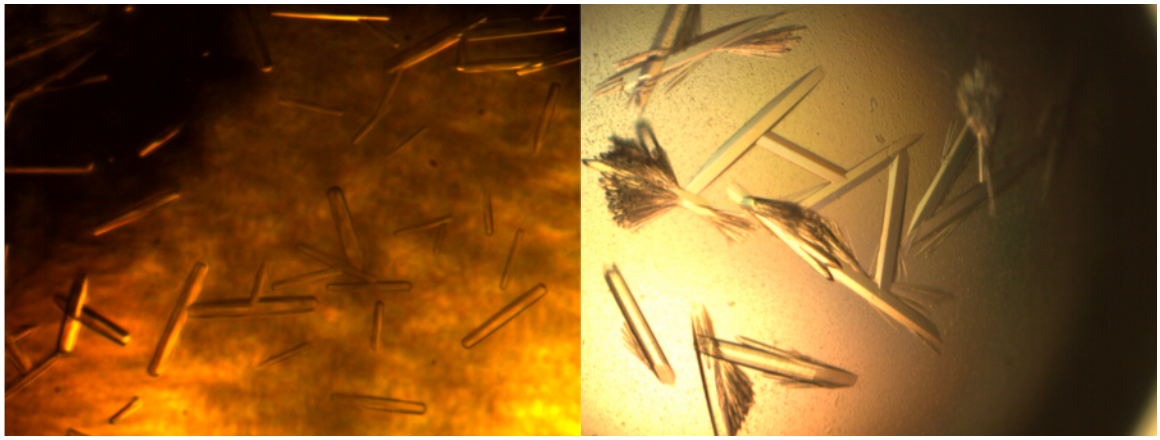


Figure 4.7: Preparative gels of RioK1 (143-494) purification. Expected molecular weight is 40.89 kDa. Lane M is the protein ladder **A.** Gel after first HisTrap column purification, lane 1 is the lysate and lane 2 is the flow-through after loading the protein onto the column. **B.** Gel after the HisTrap nickel column purification, lane 1 is before histidine tag cleavage and lane 2 is after cleavage. **C.** Gel after size exclusion chromatography.

A. **B.**



C.

Figure 4.8: RioK1 (143-494) crystals. **A.** Crystals observed in Wizard III screen before optimization. **B.** Crystals after optimization in 28% PEG 400, 0.1 M HEPES pH 7.5 and 0.2 M $CaCl_2$. **C.** Se-Met RioK1 (143-494) optimized crystals that grew in 28% PEG 400, 0.1M HEPES pH 7.5 and 0.2 M $CaCl_2$.

Table 4.3: Scalepack log file: Se-Met RioK1 (143-494)

Shell Lower limit	Shell Upper limit	Average Redundancy Per Shell
50.00	5.66	4.7
5.66	4.50	4.8
4.50	3.93	4.8
3.93	3.57	4.8
3.57	3.31	4.8
3.31	3.12	4.8
3.12	2.96	4.7
2.96	2.83	4.1
2.83	2.72	3.6
2.72	2.63	3.5
All hkl		4.5

Shell Lower limit	Shell Upper limit	I/Sigma in resolution shells: % of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
50.00	5.66	0.1	0.2	0.4	0.8	1.5	9.6	99.5	0.0	99.5
5.66	4.50	0.1	0.6	1.4	2.4	4.9	16.4	100.0	0.0	100.0
4.50	3.93	0.3	1.8	3.7	5.0	10.9	24.3	99.0	1.0	100.0
3.93	3.57	1.5	4.0	8.2	13.3	22.7	41.5	76.9	23.1	100.0
3.57	3.31	3.1	10.1	18.9	28.8	41.9	61.3	80.0	20.0	100.0
3.31	3.12	9.2	20.2	33.6	44.6	58.0	73.4	85.5	14.5	100.0
3.12	2.96	14.3	32.6	49.7	59.2	68.7	80.6	90.0	10.0	100.0
2.96	2.83	18.4	41.6	58.3	66.9	75.5	84.5	92.7	6.5	99.2
2.83	2.72	19.6	44.1	59.5	65.7	70.4	77.3	83.4	5.6	89.0
2.72	2.63	15.5	35.0	46.4	51.6	57.0	61.6	67.3	4.1	71.4
All hkl		8.2	19.0	28.0	33.8	41.1	53.0	87.5	8.5	95.9

Shell limit	Lower Angstrom	Upper	Average I	Average error	Norm. stat.	Linear Chi**2	Square R-fac	R-fac
50.00	5.66		405.4	28.1	5.5	1.657	0.092	0.104
5.66	4.50		176.0	12.1	3.1	1.312	0.103	0.114
4.50	3.93		145.7	9.1	3.1	1.243	0.097	0.098
3.93	3.57		95.1	4.9	2.8	1.227	0.091	0.077
3.57	3.31		55.7	3.2	2.6	1.304	0.111	0.074
3.31	3.12		25.6	2.6	2.4	1.246	0.189	0.127
3.12	2.96		21.2	2.7	2.5	1.173	0.203	0.104
2.96	2.83		12.3	2.8	2.6	1.005	0.261	0.152
2.83	2.72		10.2	3.1	2.9	0.984	0.259	0.165
2.72	2.63		9.6	3.1	2.9	0.999	0.253	0.167
All reflections			99.1	7.3	3.0	1.241	0.106	0.102

Table 4.4: Scalepack log file: RioK1 (143-494)

Shell Lower limit	Shell Upper limit	Average Redundancy Per Shell
50.00	6.14	5.4
6.14	4.87	5.8
4.87	4.26	5.9
4.26	3.87	6.0
3.87	3.59	6.0
3.59	3.38	6.0
3.38	3.21	6.1
3.21	3.07	5.7
3.07	2.95	5.3
2.95	2.85	5.0
All hkl		5.7

Shell Lower limit	Shell Upper limit	I/Sigma in resolution shells: % of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
50.00	6.14	1.0	2.2	2.6	2.9	4.0	13.1	54.2	44.2	98.4
6.14	4.87	1.1	2.1	4.3	6.0	9.9	23.7	64.0	35.9	99.9
4.87	4.26	2.3	5.3	8.9	12.2	19.0	33.1	63.1	36.9	100.0
4.26	3.87	4.6	8.9	15.1	21.6	33.2	48.9	74.6	25.4	100.0
3.87	3.59	8.1	18.9	29.6	37.7	52.2	70.1	86.4	13.6	100.0
3.59	3.38	15.1	32.1	47.4	58.4	68.8	79.6	90.3	9.7	100.0
3.38	3.21	22.0	44.0	60.1	69.0	78.1	86.1	93.8	6.0	99.8
3.21	3.07	28.0	51.4	68.0	74.9	81.4	88.4	95.3	2.6	97.9
3.07	2.95	27.4	51.6	64.9	72.3	77.2	83.5	88.3	3.1	91.4
2.95	2.85	26.9	48.5	61.0	67.1	72.0	77.7	81.3	1.5	82.8
All hkl		13.4	26.0	35.5	41.5	48.8	59.7	78.8	18.4	97.1

Shell limit	Lower Angstrom	Upper	Average I	Average error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac
50.00	6.14		714.6	34.6	11.6	1.179	0.084	0.107
6.14	4.87		186.8	9.1	3.9	1.198	0.083	0.094
4.87	4.26		199.2	8.7	4.2	1.083	0.072	0.074
4.26	3.87		144.2	6.9	4.2	1.152	0.077	0.062
3.87	3.59		86.6	5.5	4.1	1.174	0.111	0.087
3.59	3.38		48.3	4.8	4.1	1.435	0.170	0.098
3.38	3.21		25.1	4.4	4.0	1.423	0.283	0.166
3.21	3.07		14.2	4.5	4.3	1.431	0.435	0.243
3.07	2.95		16.0	5.0	4.8	1.548	0.362	0.187
2.95	2.85		8.8	5.3	5.2	1.346	0.543	0.347
All reflections			153.9	9.2	5.1	1.288	0.102	0.100

Table 4.5: Data Collection and Refinement statistics.

RioK1 (143-494)	
Data collection	
Space group	$P3_2$
Cell dimensions	
a, *b*, *c* (Å)	78.81, 78.81, 110.57
α, β, γ (°)	90, 90, 120
Molecules/Asym. unit	2
Wavelength (Å)	0.97919
Resolution (Å)	2.63
R_{sym} (last shell)	10.6 (25.3)
$I/\sigma I$	13.6 (3.1)
Completeness (%)	95.9 (71.4)
Redundancy	4.5 (3.5)
Refinement	
Resolution (Å)	68.25-2.63
R_{work}/R_{free} (%)	22.8/26.6
Mean B-factors $(Å)^2$	42.96
RMS deviations	
Bond lengths (Å)	0.0168
Bond angles (°)	1.8768
Ramachandran plot	
Favored	94.6
Additional Allowed	5.4
Disallowed	0.0

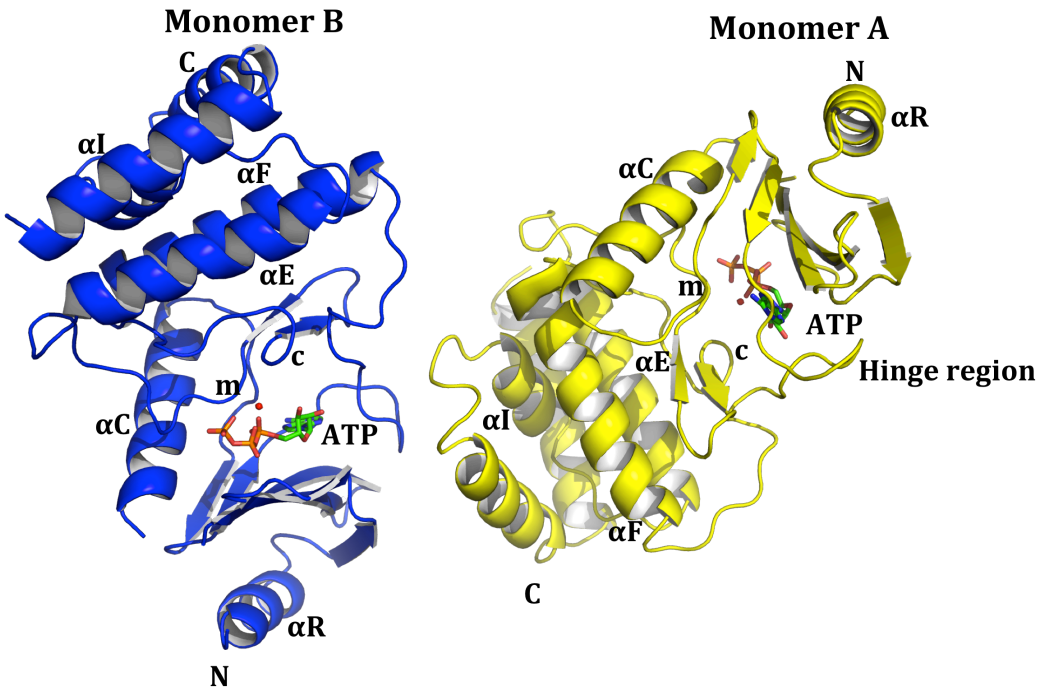


Figure 4.9: **RioK1 (143-494) overall structure**. RioK1 (143-494) overall structure characterized by a β-sheet on the N-lobe along with a sub-family specific helix, αR, and a α-helical C-lobe. The helices are labeled according to afRio1 structure and the metal binding loop and catalytic loop are labeled as m and c respectively.

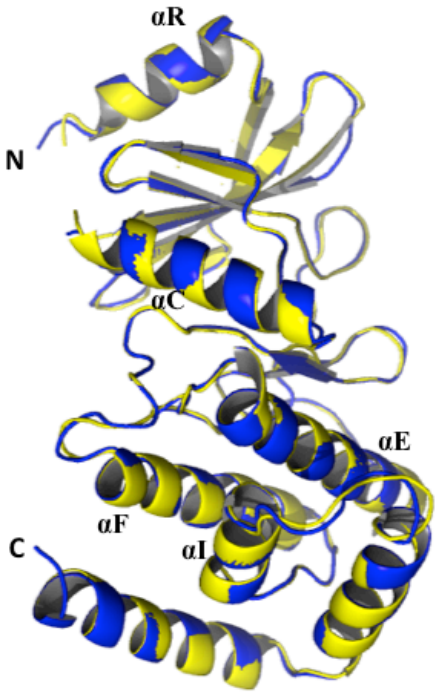


Figure 4.10: RioK1 (143-494)-ATP/Mg^{2+} alignment of the two molecules per asymmetric unit. Monomer A is in blue and monomer B in yellow.

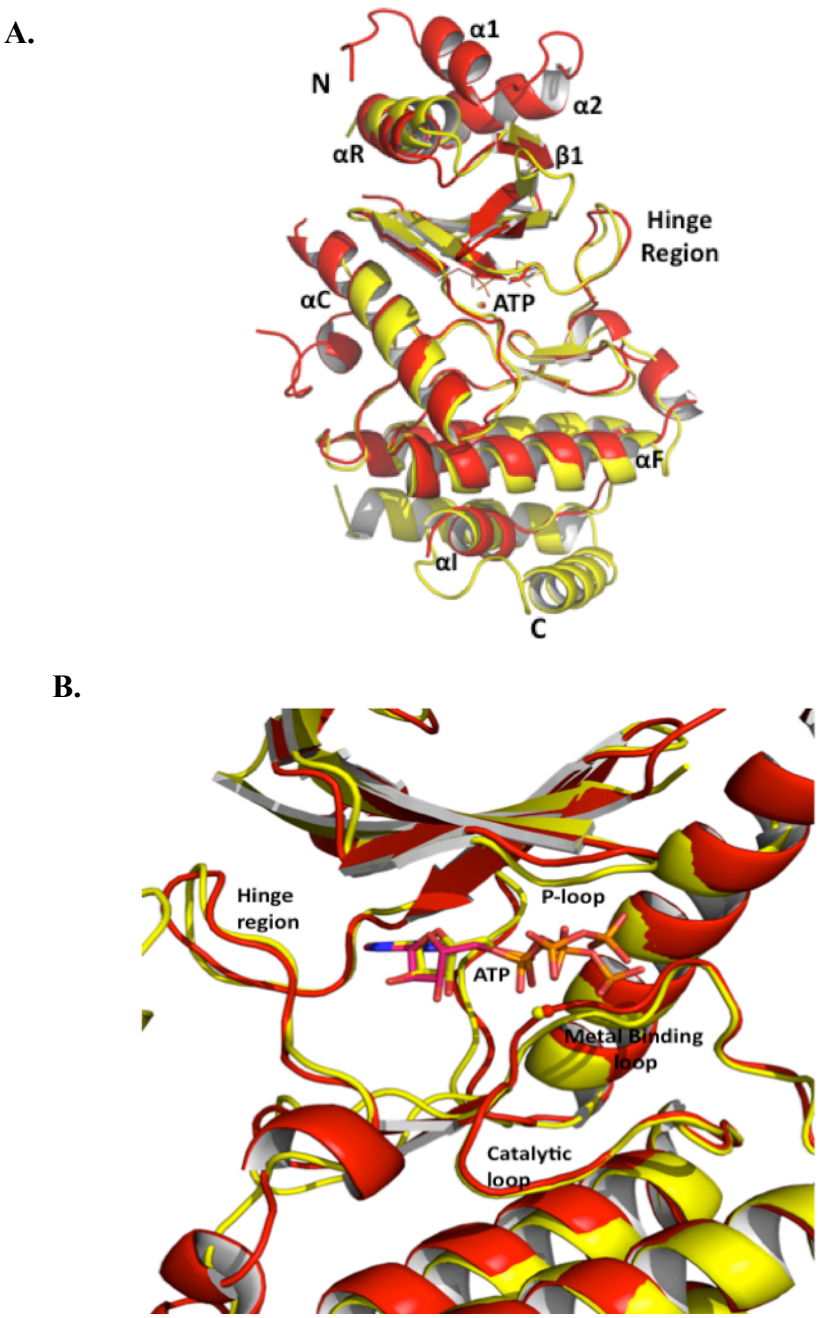


Figure 4.11: Alignment of RioK1 and afRio1. RioK1 structure is in yellow and afRio1 is red. **A.** Overall alignment of RioK1 and afRio1 (PDB ID: IZP9). **B.** Alignment of the active site of RioK1 and afRio1. Shifts are observed in the N-lobe and the metal-binding loop.

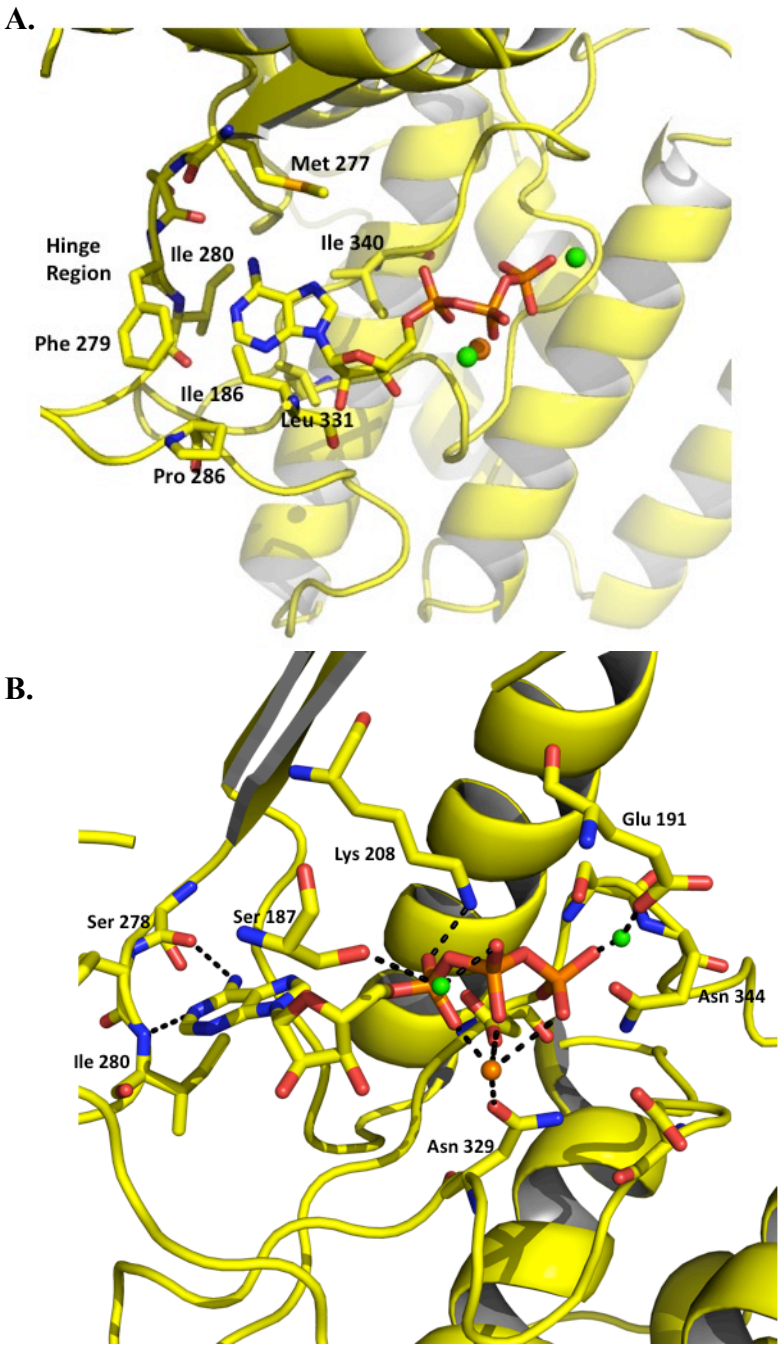
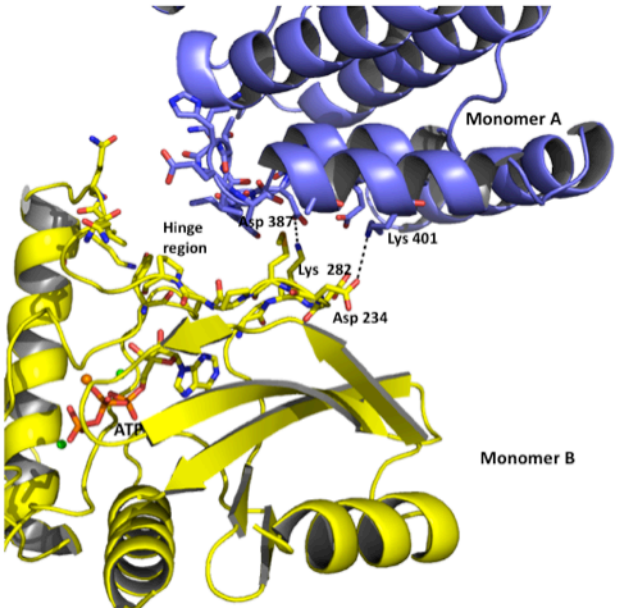


Figure 4.12: RioK1 (143-494)-ATP/Mg^{2+} interactions. **A**. Shows hydrophobic interactions. **B.** Shows hydrogen bonds as black dashed lines. Spheres in green are waters and the orange sphere is $Mg^{2+.}$

A.



B.

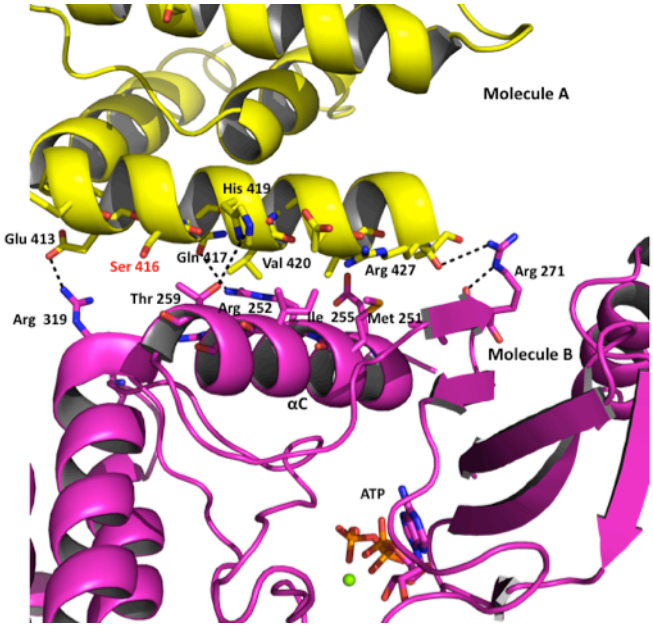


Figure 4.13: RioK1 potential interacting interfaces. **A.** Interface between two molecules in the asymmetric unit. Monomer A is in blue and monomer B is in yellow. **B.** Interface between symmetry related molecules, molecule A is in yellow and molecule B is in magenta. The autophosphorylation site, ser 416 (in red) is found in this interface between the symmetry related molecules.

Table 4.6: Primers for RioK1 mutants

Mutants	**Primer sequence**
RioK1 (106-494)	5′-AAATAGCATCTCAAAGGACCGCGGCAGAACGGTCTAGCCAAGATC-3′
K269AE270A	5′-GATCTTGGCTAGACCGTTCTGCCGCGGTCCTTTGAGATGCTATTT-3′
RioK1 (106-494)	5′-AAATTAATTTAGATAAGCTAAATGTTACTGATGCCGTCATAAATAAAGTCACCG-3′
S130A	5′-CGGTGACTTTATTTATGACGGCATCAGTAACATTTAGCTTATCTAAATTAATTT-3′
RioK1 (106-494)	5′-GGACCAAGGAAGAACGGGCTAGCCAAGATCATGTG-3′
S415A	5′-CACATGATCTTGGCTAGCCCGTTCTTCCTTGGTCC-3′
RioK1 (106-494)	5′-GGACCAAGGAAGAACGGTCTGCCCAAGATCATGTGG-3′
S416A	5′-CCACATGATCTTGGGCAGACCGTTCTTCCTTGGTCC-3′

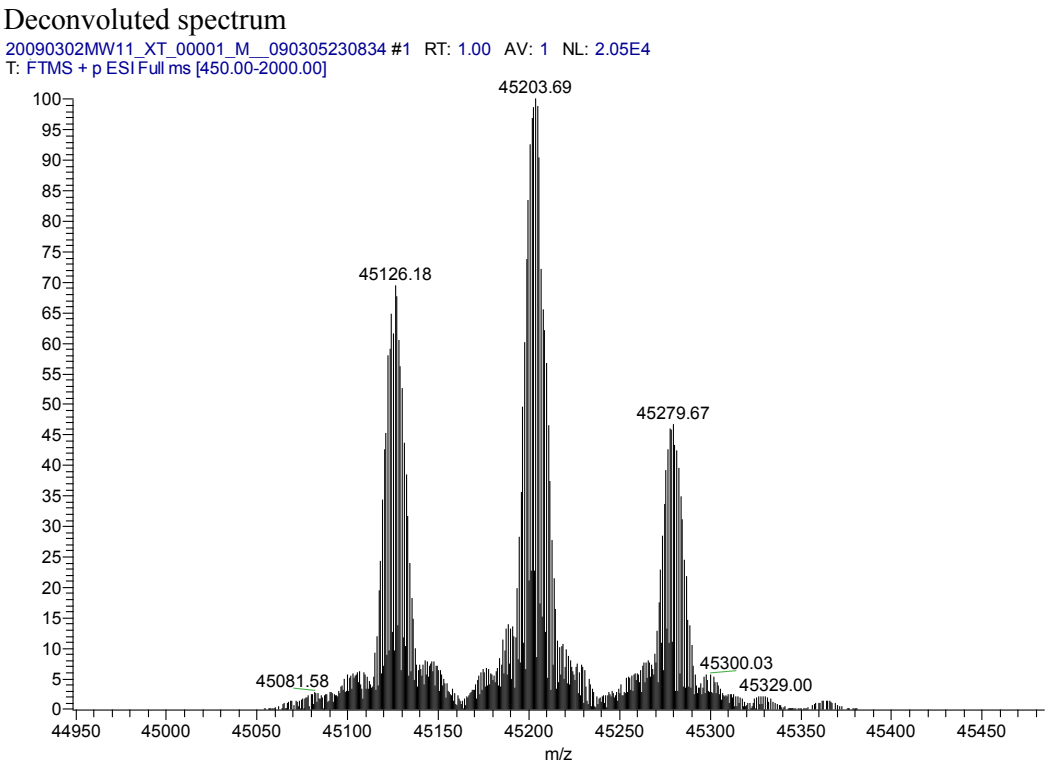


Figure 4.14: RioK1 (106-494) Electrospray Ionization Spectrum. Three peaks are observed corresponding to the expected molecular weight, 45126.18 Daltons, singly phosphorylated RioK1 (106-494), 45203.69 Daltons and doubly phosphorylated RioK1 (106-494), 45279.67 Daltons.

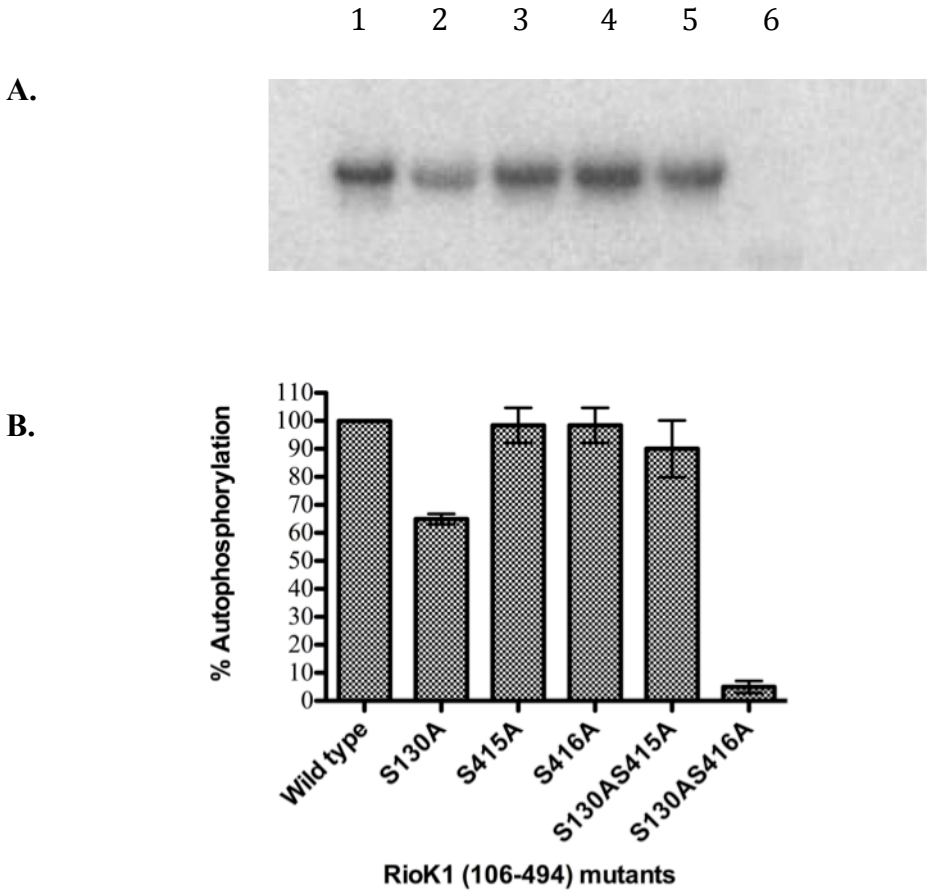


Figure 4.15: Identifying RioK1 autophosphorylation sites. **A**. Denaturing gel showing autophosphorylation activity. Lane 1 is wild-type protein, lane 2 is the mutant S130A, lane 3 is the mutant S415A, lane 4 is the mutant S416A, lane 5 is the mutant S130AS415A and lane 6 is S130AS416A. **B.** Graph representing percent of autophosphorylation activity. From both the gel and graph it's clear that the two sites are Ser 130 and Ser 416.

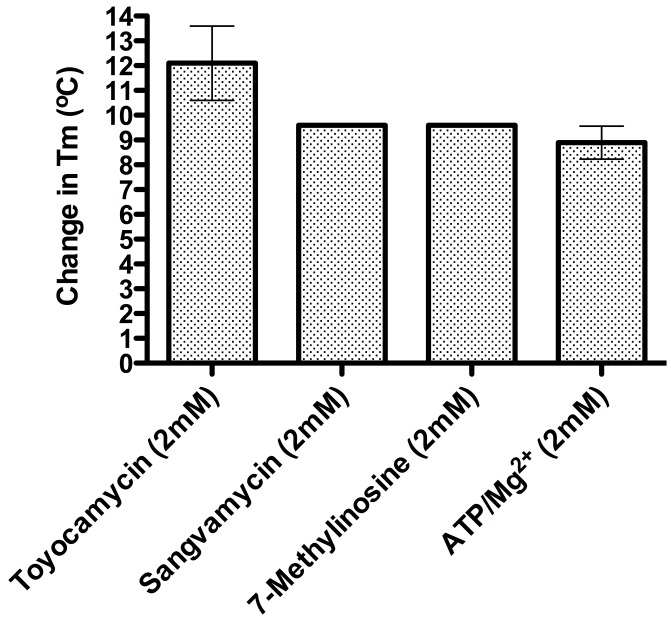


Figure 4.16: RioK1 Thermoflour assay. Toyocamycin has the biggest T_m shift of 12.1°C comparable to afRio1 studies. Toyocamycin binds with a higher affinity and stabilizes RioK1 more than ATP/Mg^{2+}.

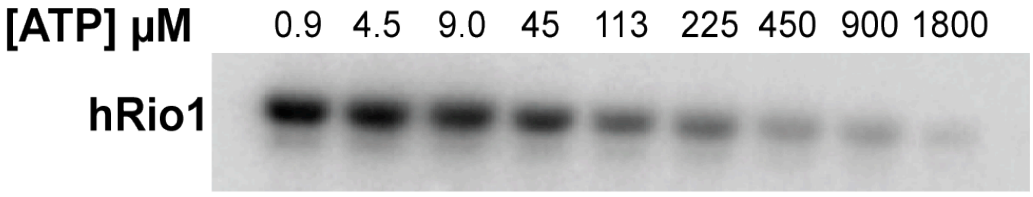
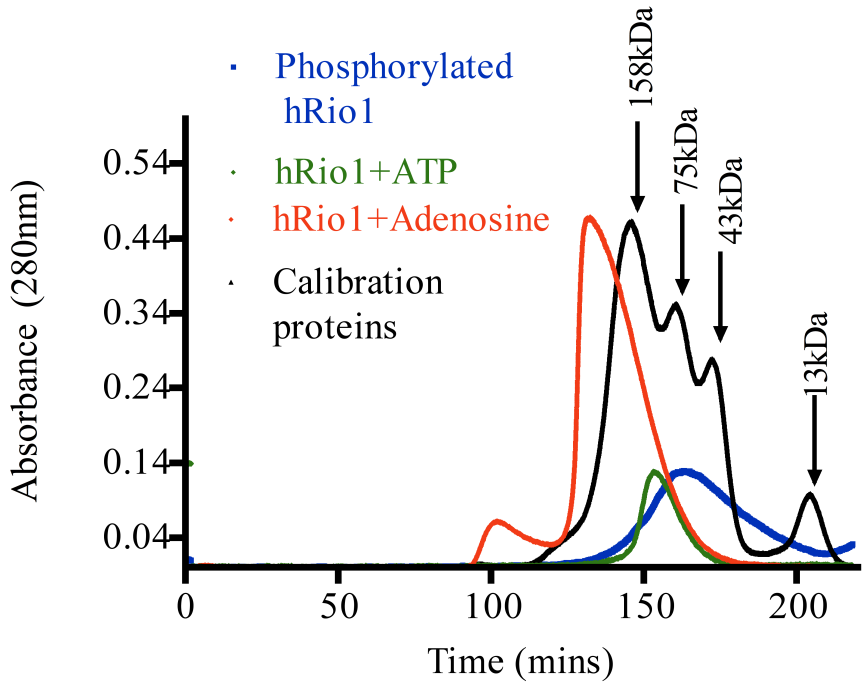


Figure 4.17: Substrate inhibition assay. At high concentrations of ATP, RioK1 autophosphorylation activity is diminished whereas at low concentrations of ATP such as 0.9 µM, activity is observed.

A.



B.

Sample	**Calculated MW (kDa)**	**Oligomer**
Phosphorylated RioK1	70.3	Monomer-Dimer
RioK1+ATP	104.3	Dimer
RioK1+Adenosine	179.1	Tetramer

Figure 4.18: RioK1 size exclusion chromatography data. Oligomers are observed in the presence of adenosine and ATP. **A.** Shows the size exclusion chromatogram. **B.** Table with calculated molecular weight and corresponding oligomeric state.

Chapter 5: Functional Studies of afRio1 Kinase

5.1 Overview

Crystal packing analysis of afRio1 structure in the absence of ATP revealed contact between two symmetry-related molecules via the flexible loop of one molecule and the active site of the second molecule (Figure 5.1). This interaction positions the autophosphorylation site, serine 108, of one molecule close to the active site of the second molecule and involves several conserved residues that do not appear to be involved in ATP binding based on the crystal structure of afRio1 with ATP bound (Figure 5.1). These residues are conserved among Rio proteins and across species suggesting their involvement in a significant role, possibly peptide or protein substrate binding, recognition and interaction. This chapter focuses on determining how conserved residues found near the active site affect afRio1 autophosphorylation and kinase activity. The following residues Lys 59, Glu 60, Arg 83, and Tyr 200 were mutated and tested for activity.

5.2 Experimental

5.2 a Mutagenesis, expression and purification of afRio1 mutants.

Site-directed mutagenesis was carried out using QuikChange II kit (Agilent) according to manufacturer's instructions. AfRio1 mutants, K59A, E60A, R83A, and Y200D were expressed and purified following the same protocol as wild-type afRio1 with three modifications. After lysis, the lysate was heated at 68°C instead of 75°C for 15 minutes to denature *E.coli* proteins, the N-terminal histidine tag was not

cleaved and ion exchange was the final purification step instead of size exclusion chromatography.

After the first HisTrap column purification, all mutants were dialyzed against a buffer containing 50 mM Tris pH 8.0, 200 mM NaCl and 0.2% β-mercaptoethanol followed by dilution with different buffer containing 50 mM Tris pH 8.0 and 0.2% β-mercaptoethanol for a final salt concentration of 100 mM NaCl. All mutants were loaded onto an ion exchange column, the Q column (GE healthcare) at rate of 1.0 mL/min and subjected to a salt gradient of 100 mM NaCl to 1 M NaCl. AfRio1 mutants eluted between 200 mM and 300 mM NaCl and all purification steps were monitored by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE gel) for purity (Figure 5.2).

5.2b Autophosphorylation and Kinase activity assays.

For autophosphorylation assays, 5 μg of purified mutants, K59A, E60A, R83A, Y200D and wild-type afRio1 were added to a 30 μl reaction containing 50 mM Tris pH 8.0, 50 mM NaCl and 2 mM $MgCl_2$. 5 μCi of γ-^{32}P labeled ATP was added to the reactions and incubated at 37°C for 1 hour. For assays used to test kinase activity, 0.5 μg instead of 5 μg of purified protein was added to the reactions along with 10 μg of myelin basic protein (MBP), a general kinase substrate. All reactions were carried out in triplicates, stopped by adding 8 μl of SDS loading dye to denature the protein. The reaction mixtures were ran on NuPAGE®4-12% Bis-Tris denaturing gels (Invitrogen) at 150V for 45 minutes to separate the protein from γ-^{32}P labeled ATP. Gels were dried and exposed on a phosphoimager screen for 15 hours followed

by quantification using the software ImageQuant (Molecular Dimensions) to determine autophosphorylation and kinase activity (Figure 5.3).

5.3: Results

AfRio1 E60A mutant had the least amount of autophosphorylation and kinase activity whereas afRio1 K59A had the most autophosphorylation activity but a diminished kinase activity, second to afRio1 E60A (Figure 5.3 and 5.4). Glu 60 and Lys 59 are both found on the nucleotide-binding loop (p-loop) of Rio kinases and though they are found near the active site, they do not participate in binding interactions with ATP. The autophosphorylation results suggest that Lys 59 is not vital in phosphorylating afRio1 since its mutation has no effect on autphosphorylation activity. However, Lys 59 seems to be important in phosphorylating protein substrates such as MBP as per kinase assays. On the other hand, the results clearly indicate the importance of Glu 60 in both autophosphorylation and phosphorylation of protein substrates.

AfRio1 Y200D mutant had both diminished autophosphorylaytion and kinase activities with more effect on autophosphorylation (Figure 5.3 and 5.4). Tyrosine 200 that is found on the catalytic loop is either a tyrosine or a phenylalanine in RIO kinases. AfRio1 Y200D's diminished catalytic activity is not surprising for a mutation on a residue that is found on the catalytic loop. Finally Arg 83, that is either an arginine or a lysine in Rio1 kinases, does not affect autophosphorylation or kinase activity to a large extent, indicating that the residue, though conserved, might not be important for protein substrate recognition, interactions or activity.

5.4 Summary

This chapter identifies conserved residues, Lys 59, Glu 60 and Y200 to be vital for afRio1 activity. Though the residues are not involved in ATP binding, they do seem to have another vital role, possibly protein or peptide substrate binding, recognition and interaction. Experiments with Rio1-specific substrates, which have not yet been identified, as well as crystal structures of a mutant-substrate complex would aid in identifying with certainty the role of these residues. It would also be of interest to see how the mutants affect the oligomeric state of afRio1, which in solution is a mixture of monomer-tetramer (Chapter 3).

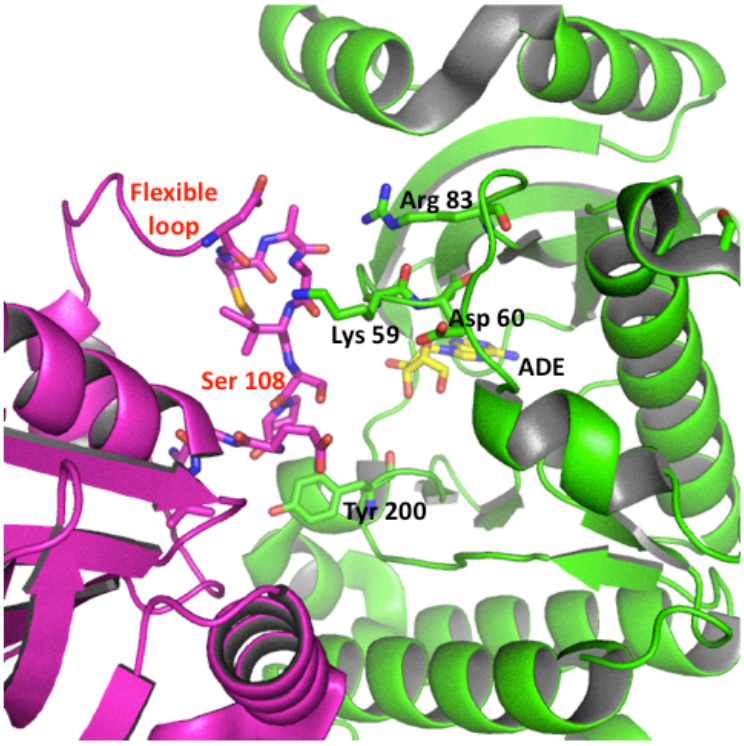


Figure 5.1: Interface between two symmetry-related molecules of afRio1 without ATP bound. Contact is between the flexible loop of one molecule (magenta) and the active site of the second molecule (green). Conserved residues near the active site are that were mutated are labeled on one molecule (green). Residual adenosine, labeled Ade, is observed in the active site.

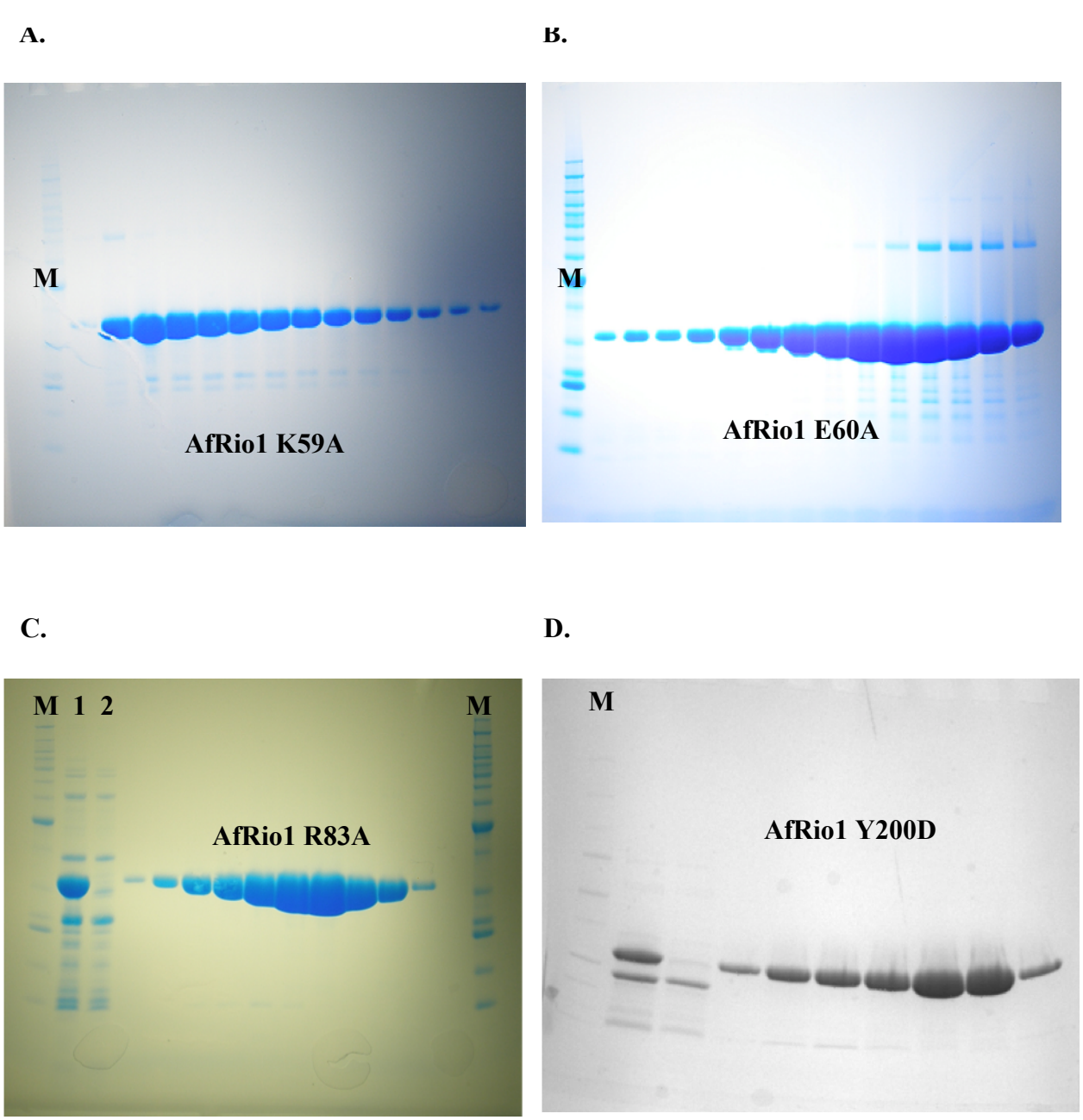


Figure 5.2: **SDS preparative gels of afRio1 mutants after ion exchange purification**. Lane M in all gels is the protein ladder. **A.** AfRio1 K59A mutant. **B.** AfRio1 E60A mutant. **C.** AfRio1 R83A mutant, Lane 1 is the lysate sample and Lane 2 is uninduced sample. **D.** AfRio1 Y200D mutant.

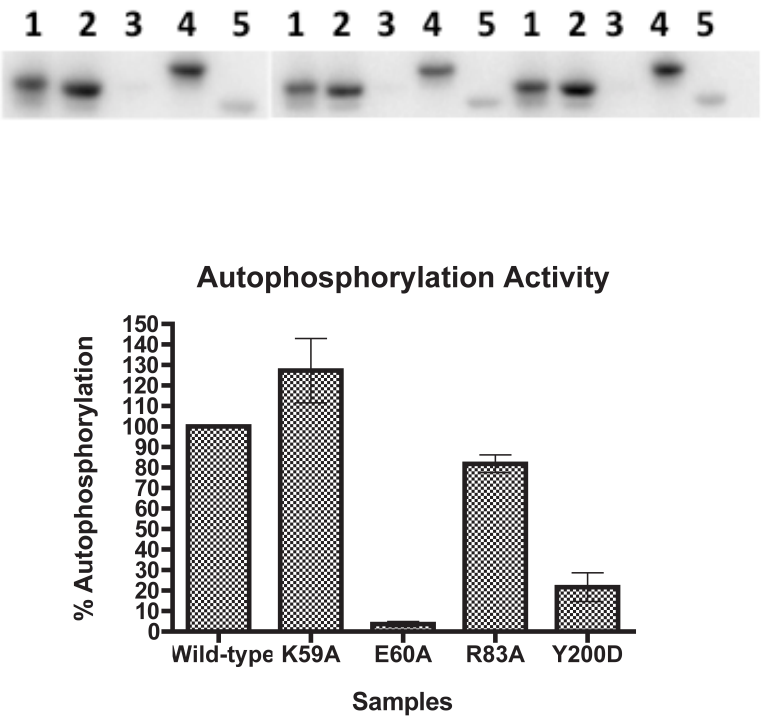


Figure 5.3: Autophosphorylation assay. A. Preparative gel showing autophosphorylation activity. All reactions were done in triplicates; E60A mutant has the least activity whereas K59A has the most activity. Lane 1 is wild-type afRio1 reaction, lane 2 is K59A reaction, lane 3 is E60A reaction, lane 4 is R83A reaction and lane 5 is Y200D reaction. **B.** Graph showing percent of autophosphorylation activity.

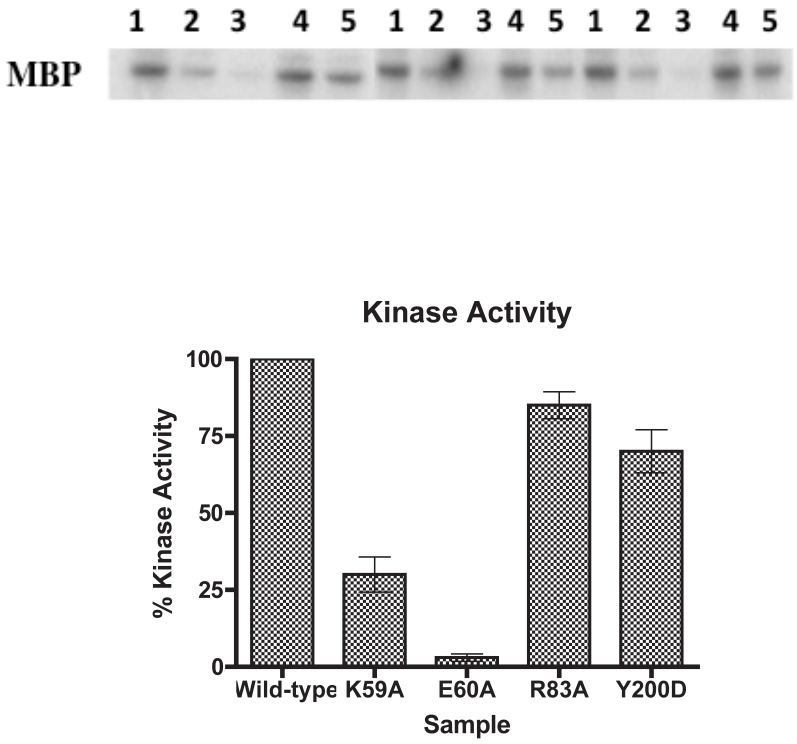


Figure 5.4: **Kinase assay used to determine activity by amount of myelin basic protein (MBP) phosphorylation**. **A.** Preparative gel showing phosphorylation of MBP. All reactions were done in triplicates; E60A mutant has the least activity whereas the wild-type has the most activity. Lane 1 is wild-type afRio1 reaction, lane 2 is K59A reaction, lane 3 is E60A reaction, lane 4 is R83A reaction and lane 5 is Y200D reaction. **B.** Graph showing percent of kinase activity.

Chapter 6: Structural determination of EspB

6.1 Introduction

Tuberculosis (TB) is an airborne infectious disease caused by *Mycobacterium tuberculosis (Mtb)* that affects the respiratory system. In 1993, the World Health Organization (WHO) declared tuberculosis to be a global health emergence and remains to be one of the world's major causes of illness and death (*101*). A third of the world's population is infected with *M. tuberculosis* and most infections affect people in resource-poor settings. For instance 90% of new TB cases and deaths occur in developing countries due to poor health systems, treatment barriers (unreliable drug supply or patients not completing treatment) and limited resources for disease detection. In addition, the emergence of drug resistant TB as well as TB and HIV co-infections pose a threat to global control of the disease (*102*). In 2008, of the estimated 1.8 million people who died from TB, 500,000 were HIV-positive making it the leading cause of death among people living with HIV (*102*). TB treatment can also take a long time and be expensive especially due to drug-resistant TB of which there are two forms, multidrug-resistant TB (MDR-TB) and extensively drug-resistant TB (XDR-TB) (*103*). MDR-TB is resistant to standard first line drugs such as rifampicin and XDR-TB is resistant to both first and second line drugs such as amikacin(*102*) (*104*). The severity of this disease necessitates a better understanding of *Mtb's* pathogenic mechanism with the aim of developing new therapeutics to combat TB.

Mycobacterium tuberculosis (Mtb) was first discovered in 1882 by Robert Koch and is one of the 85 species in the mycobacterium genus. *Mtb* has a thick cell

wall that is covalently linked by a layer of branched-chain fatty acids known as mycolic acids. This provides a protective, hydrophobic waxy coat known as the mycomembrane. *Mtb* requires high levels of oxygen and thrives well in the human lungs. Upon *Mtb* infection, a phagosome is formed via phagocytosis of the bacterium by macrophages. The phagosome undergoes maturation by first fusing with various endosomes via the endocytic pathway followed by fusion with lysosomes to form phagolysosomes (*105-107*). During phagolysosome formation, degradative enzymes that destroy invading pathogens are released and the pH drops to about 4.5, an environment that is not conducive for *Mtb* survival(*108*). Phagolysosome formation is one of the multiple antimicrobial responses that the host cell employs to limit the growth of infection (*109*).

Survival of *Mtb* in macrophages is critical to the virulence and pathogenesis of the bacterium. One survival response that *Mtb* utilizes is inhibition of phagosome maturation by preventing acidification of the phagosome by lysosomes that have a pH of about 4.5. The phagosome is arrested in the endocytic pathway under a conducive pH of about 6.4 and phagolysosomes are not formed (*110*)*106*). The phagosome inhibition mechanism is not well understood but studies do show the involvement of a specialized secretion pathway known as ESX-1(*111-113*). This secretion pathway aids *Mtb* manipulate its environment by delivering virulence factors during infection.

The ESX-1 secretion system also known as Type VII secretion system, has several virulence proteins that are substrates including culture filtrate protein-10 (CFP-10), early secretory antigenic target-6 (ESAT-6) and ESX-1 substrate protein B (EspB) among others (*113, 114*). The ESX-1 secretion system is encoded for by a

chromosomal region RD1 (Region of Difference-1) comprised of nine genes (Rv3871-Rv3879c) and its surrounding region, together known as extRD1 (extended region of difference-1) (*111, 113, 115, 116*). RD1 was identified via comparative studies of complete genomes of various mycobacterial species. For instance the less virulent species, *Mycobacterium bovis* Bacille Calmette-Guérin (BCG) lacks this chromosomal region that is present in more virulent species such as *Mtb* and *M. marinum (117, 118). M. marinum* is genetically closely related to *Mtb* and causes TB-like diseases in its natural hosts (fish and frogs). *M. marinum* is therefore used in understanding *Mtb* pathogenesis. ESX-1 also has core components that include ATPase members (Rv3870, Rv 3871, Rv 3876), a transmembrane protein (Rv 3877) and MycP1, a substilin-like serine protease (*119*) (Figure 6.1).

ESX-1 secretion has been implicated in inhibition of phagosome maturation, escape of *Mtb* from endocytic membrane, apoptosis of macrophages, and formation of pores in mycobaterial phagosomes allowing for the spread of bacteria (*115, 120-123*). ESX-1 is unique in that substrate secretion is mutually dependent (*124*). For instance it has been shown that secretion of CFP-10, ESAT-6 and EspB is co-dependent and ESAT-6 interacts with the C-terminus of EspB(*113*). Studies also show that CFP-10 and ESAT-6 form a 1:1 heterodimer complex that interacts with the ATPase component Rv3871 (*114*). In addition, the N-terminus of EspB interacts with the C-terminus of Rv3879c, which then interacts with Rv3871 delivering EspB to the Rv3871-CFP-10/ESAT-6 secretory machine (*114*). In this role, Rv3879c is hypothesized to be a cytosolic chaperone and these interactions are necessary for secretion. Once EspB is secreted to the periplasmic space, the protease MycP1 that is

localized near the cell membrane cleaves EspB at two sites, Ala 392 and Pro 332 on the C-terminus (*119*). EspB has a third cleavage site but the protease responsible has not yet been identified. While EspB's C-terminus is necessary for interaction and secretion of CFP-10/ESAT-6 complex, it is not required for EspB secretion. This is supported by studies that show an EspB fragment with the C-terminus truncated is secreted by ESX-1(*113, 114*). Further evidence for EspB's C-terminus involvement in secretion is found in a system that has a MycP1 mutant (MycP1 S332A) whose protease activity is abolished. This mutant is not able to cleave EspB and consequently ESX-1 secretion is upregulated as EspB's C-terminus is available for interaction and secretion of ESX-1 substrates (*119*).

ESX-1 secretion pathway is not fully understood and some unknowns that interest my current lab include: a) the molecular structure of EspB

b) EspB oligomeric state and how this relates to different cellular environments that *Mtb* encounters, c) interaction of EspB with ESAT-6 and Rv3879c, and d) the role of EspB proteolytic fragments, their oligomeric states, their interactions and functions in different cellular compartments. This chapter describes the work that went towards structural determination of EspB and presents the first crystal structure of EspB along with structural analysis.

6.2 Experimental

6.2a Expression of Se-Met EspB 1-292 and EspB 1-346

Two constructs of EspB from *M.Marinum*, EspB residues 1-292 and EspB residues 1-346 were expressed, purified and crystallized. These constructs were designed based on previous work that had shown degradation of full-length EspB

during purification, mass spectrometry analysis by a former collaborator, Dr. Gao, showing cleavage at Leu 346 and limited proteolysis results that identified EspB (1-292) as a stable fragment. EspB (1-292) and EspB (1-346) plasmids with N-terminal histidine tags were transformed to Rosetta™(DE3) pLysS *E.coli* cells (Novagen). EspB (1-292) was expressed in 1 L of seleno-methionine minimal media (Shanghai, Medilon Inc) with 34 μg/mL chloramphenicol and 100 μg/mL ampicillin, to generate Se-Met protein. The cells were grown at 37°C while shaking at 250 rpm. Once the OD_{600} reached 1.2, 1 mM of Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the media for induction and expression proceeded at 18°C for 24 hours. The cells were centrifuged at 5,000 rpm and 4°C for 20 minutes using a Sorvall centrifuge (Dupont) with a SLC 4000 rotor. The bacterial pellet was stored at -80°C. EspB (1-346) was expressed in 2 L of LB media following the same expression procedure as EspB (1-292). The strategy was to solve the structure of EspB (1-292) by single anomalous dispersion (SAD) phasing and then use this model for molecular replacement to solve EspB (1-346).

6.2b Purification of Se-Met EspB (1-292) and EspB (1-346)

The frozen bacterial cell pellets were resuspended in lysis buffer containing 50 mM Tris pH 8.0, 200 mM NaCl, 0.05% β-mercaptoethanol, 10% glycerol and one tablet of protease inhibitor. The resuspended bacterial cell pellets were lyzed using a French press and the lysate was centrifuged at 4°C and 18,000 rpm for 40 minutes using a Beckman Coulter ultracentrifuge with a 45 Ti rotor. The supernatant was filtered and imidazole added for a final concentration of 20 mM. This supernatant was

loaded onto a pre-equilibrated 5mL HisTrap™ HP column (GE healthcare) attached to a Fast Protein Liquid Chromatography (FPLC) system at a flowrate of 2.0 mL/min and washed with 10 column volumes of lysis buffer to remove any unbound protein. Bound protein was subjected to an imidazole gradient of 20 mM to 250 mM and Se-Met EspB (1-292) eluted between 100 mM and 150 mM imidazole whereas EspB (1-346) eluted between 120 mM and 160 mM imidazole. Partially purified fractions were pooled together and transferred to a dialysis membrane to which 1 mg/mL of Tobacco etch virus (TEV) protease was added to cleave the histidine tag. Se-Met EspB (1-292) was dialyzed overnight at 4°C against 20 mM Tris pH 8.0, 50 mM NaCl, 0.05% β-mercaptoethanol and 10% glycerol whereas EspB (1-346) was dialyzed against the lysis buffer. After cleavage both proteins were loaded onto a pre-equilibrated 5mL HisTrap™ HP column (GE healthcare) and Se-Met EspB (1-292) washed with the same buffer used for dialysis whereas EspB (1-346) was washed with the lysis buffer. Both cleaved proteins were further purified by size exclusion chromatography using a Superdex™ 200 column (GE healthcare) equilibrated in 10 mM HEPES, 50 mM NaCl and 0.05% β-mercaptoethanol at pH 7.5 for Se-Met EspB (1-292) and pH 8.0 for EspB (1-346). All purification steps were monitored by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE) for purity (Figure 6.2).

6.2c Crystallization of Se-Met EspB (1-292) and EspB (1-346)

Purified Se-Met EspB (1-292) was concentrated to 22 mg/mL and EspB (1-346) to 39 mg/mL. Both proteins were subjected to robotic sparse matrix screening

using Phoenix Liquid Handing System (Art Robbins). Commercially available matrix screens, Index (Hampton), Cyros (Qiagen), PEGs suite (Qiagen), Wizard I, II and III (Emerald Biosystems) were used for initial crystal screening in 96 well plates with three different protein to well solution ratios (1:2, 1:1, 2:1). Crystals observed during initial screening were optimized to generate diffraction quality crystals. For Se-Met (1-292), three different crystal forms, hexagonal, rectangular and diamond-shaped, were observed in Index, Wizard I and III respectively (Figure 6.3). The diamond-shaped crystals grew in 1.2 M NaH_2PO_4, 0.8 M K_2HPO_4, 0.1 M CAPS pH 10.5, 0.2 M Li_2SO_4 and optimized crystals were observed with pH 8.5 and 9.5. The rectangular-shaped crystals grew in 0.2 M $MgCl_2.2H_2O$, 25% PEG 3350, 0.1 M Tris pH 8.5 and after optimization additional crystals were observed in pH 6.5 and 7.5. Finally the hexagonal-shaped crystals grew in 20% PEG 3350, 0.2 M Magnesium formate and after optimization crystals were observed in pH 6.5, 7.5, 8.5 and 9.5. Diffraction quality crystals were flash frozen in their respective well conditions with 20 % (v/v) ethylene glycol as the cryoprotectant. It was clear that crystals grew in different pH conditions, including pH 6.5 that mimics the intracellular pH of a phagosome. Consequently an attempt was made to determine the crystal structure of different crystal forms as well as crystals from different pH conditions. The goal was to determine whether different pH ranges and crystal forms resulted in different protein conformations/overall structure.

EspB (1-346) crystals were observed in Wizard I and grew in 0.2 M NaCl, 25% PEG 3350 and 0.1 M Bis Tris pH 5.5. The rod-shaped crystals were optimized and diffraction quality crystals were observed in pH 7.5 (Figure 6.4). As with

EspB (1-292), the crystals were flash frozen with 20% (v/v) glycerol as the cryo-protectant and shipped to Argonne, IL.

6.2d Data collection and structure determination

Data was collected at 100K at the NE-CAT beamline at the Advanced Photon Source (APS), Argonne, IL, USA. The diamond-shaped crystals diffracted to 2.5 Å whereas the hexagonal-shaped crystals diffracted to 6.0 Å. Data was indexed, integrated and scaled using HKL2000 package program (*71*). The better diffracting crystals belonged to $P2_1$ space group and the unit cell dimensions were a= 173.45 Å, b= 81.70 Å, c= 267.47 Å and $\alpha=90.00^0$, $\beta=108.77^0$, $\gamma=90.00^0$. The overall redundancy was 3.7 with 99.5% data completion and an overall R_{symm} of 14.2% (Table 6.1). The structure was solved using an automated experimental phasing program, Autosol, that is part of PHENIX (*93*). Autosol found 144 heavy atoms sites using hybrid substructure search (HYSS). Experimental phases were calculated using Phaser and improved by density modification using RESOLVE (*90-92)*. After obtaining density-modified phases the program carries out a round of model building and refinement. The initial model was first subjected to two rounds of model building using AUTOBUILD that is part of PHENIX (*97-100*). Subsequent rounds of model building were carried out in COOT accompanied by several rounds of refinement using *phenix.refine* (*75, 76*). Waters were added to the structure using PHENIX and the model has a R_{work}/R_{free} of 20.6/24.6 (Table 6.2). The refined crystal structure had sixteen molecules per asymmetric unit (ASU) (Figure 6.5).

The hexagonal-shaped crystals that diffracted to 6.0 Å belonged to P3 space

group. The unit cell dimensions were a= 123.55 Å, b= 123.55 Å, c= 164.98 Å and $\alpha=90.00^0$, $\beta=90.00^0$, $\gamma=120.00^0$ with an overall redundancy of 4.6, an overall R_{symm} of 17.0% and 99.7% data completion (Table 6.4). EspB (1-346) crystals diffracted to 6.75 Å and belonged to P222 space group with cell dimensions a= 91.89 Å, b= 93.28 Å, c= 156.09 Å and $\alpha=90.00^0$, $\beta=90.00^0$, $\gamma=90.00^0$. The overall redundancy was 4.0, with an overall R_{symm} of 13.6% and 99.5% data completion (Table 6.5). Attempts were made to solve the low-resolution structures by molecular replacement using MOLREP that's part of CCP4i. Se-Met EspB (1-292) that diffracted to 2.5 Å was used as a search model (*73*).

6.3 Structural analysis of Se-Met EspB (1-292)

6.3 a Overall Structure of EspB (1-292)

EspB has no known structural or sequence homologs of known function and this chapter presents the first crystal structure of EspB. Se-Met EspB (1-292) crystal structure has 16 monomers in the asymmetric unit (molecules A-P) arranged into a set of four homotetramers. Two sets of these tetramers (tetramers 1 and 3; tetramers 2 and 4) that are arranged parallel to each other (Figure 6.5). Contact is observed between the C-terminal loop of one tetramer and the C-lobe helices of the adjacent tetramer (Figure 6.5). For instance in tetramer 2, the C-terminal loop of one monomer within that tetramer contacts the C-lobe helices of a monomer within tetramer 1, similar contacts are observed between tetramers 2 and 3 as well as 3 and 4 (Figure 6.5). One homotetramer has two monomers (monomers A and B and monomers C and D) arranged anti-parallel to each other creating an interacting interface and a

central helix bundle (Figure 6.6). Contact is also observed between the C-lobe helices of two monomers within the tetramer. For instance monomer A interacts with monomer D and monomer B interacts with C via the C-lobe helices (Figure 6.6). Each monomer is largely alpha helical (α1-α7) and is roughly shaped like a cylinder, 103.6 Å in length with a circular cross of 16.1 Å in diameter (Figure 6.7). The monomers have regions on the N- and C- terminus that lack electron density due to disorder and flexibility on these ends. All monomers, with the exception of monomer A and J, lack electron density for residues 1-8 on the N-terminus whereas eight monomers (monomer E, G, H, I, J, M, N and O) lack density for residues 287-292 on the C-terminus. In addition, all monomers lack density for residues 83-114. The overall monomeric structure also has an interesting feature on the C-terminus that is characterized by a polyproline region (residues 286-288), and of the last twenty-two C-terminal residues, ten are prolines.

Comparison of EspB (1-292) crystal structures that crystallized at different pH conditions, 8.5 and 9.5, and EspB (1-346) had the same overall structure and arrangement of tetramers. This conserved arrangement of tetramers suggests a possible biological significance.

6.3b Interfaces

The arrangement and position of monomers within a tetramer reveals two intra-tetramer interfaces, one between the anti-parallel monomers (monomer A and B) and another between the C-lobes helices of two monomers (monomer B and C) (Figure 6.8A). PDBe PISA (protein interfaces, surfaces and assemblies) server was

used to analyze these interfaces as well as calculate the buried surface area. The interface between the C-lobe helices involves α4, α5 and α6 of both monomers and has the largest calculated surface area of 1040.9 $Å^2$. This interface is characterized by hydrophobic contact between Val 150 on monomer D and Val 230 on A, Leu 217 of monomer D and Ala 227 on A, and Leu 226 on D with Leu 217 on A. Hydrogen bond interactions involving side chains are observed between; Asp 154 on D and Lys 237 on A; Arg 221 on D and Asp 228 on A; two hydrogen bonds between the side chains of Arg 220 on D and Glu 141 on A; Glu 141 on D and Arg 220 on A; Lys 237 on D and Asn 151 on A; Arg 234 on D and Gln 214 on A; Glu 233 on D and Thr 147 on A; two hydrogen bonds between side chains of Glu 231 on D and Arg 221 on A; a water-mediated interaction between Arg 221 of D and Thr 225 of A; and finally between Glu 214 on D and Arg 234 on D (Figure 6.8 A).

The second interface involving the anti-parallel monomers is characterized by interactions between the N-lobe of one monomer and a helical C-lobe (α6) on the second monomer. Interactions are also observed between an N-lobe helix (α2) on one monomer and α3 found between the N and C-lobe on the second monomer (Figure 6.8B). This interface has a calculated buried surface area of 856.8 $Å^2$ and has hydrophobic contact between Val 32 on monomer A and Phe 130 on monomer B. Hydrogen interactions are observed between the carboxylate side chain of Glu 233 on monomer A and Thr 30 on B; amine side chain of Lys 139 on monomer A and the peptidyl carbonyl oxygen of Glu 47 on B; amine side chain of Lys 134 on A and the peptidyl carbonyl oxygen of Thr 30 on B; Glu 47 on A with both Lys 139 and Lys 134 on B; Glu 47 on monomer A also has a water-mediated interaction with the

peptidyl carbonyl oxygen of Gly 126 on B; Asn 43 of A has two interactions, one with Thr 40 on B and the other with the peptidyl carbonyl oxygen of Glu 38 on B; hydroxyl group on Thr 40 on A with amine group of Asn 43 on B; and finally Thr 30 of monomer A has two interactions, one via the peptidyl carbonyl oxygen with the side chain of Lys 134 on B and another via the hydroxyl group of Thr and the carboxylate side chain of Glu 233 on B (Figure 6.8 B).

PISA analysis suggests that intra-tetramer interactions have a more significant role in complex formation and the contact between tetramers involving the C terminal of one tetramer and the C-lobe helix, α7, of the adjacent tetramer, is characterized by few residues and has a buried calculated surface area of 410 $Å^2$. Mutational studies and oligomeric state analysis would aid in identifying biological significant interfaces and effect on structure as well as function.

6.4 Conclusion

This chapter presents the first crystal structure of EspB providing objectives for biological studies such as mutational studies of residues found on intra-tetramer interfaces, oligomeric state analysis of these mutants as well as studying oligomeric state of EspB in conditions that mimic phagosome environment.

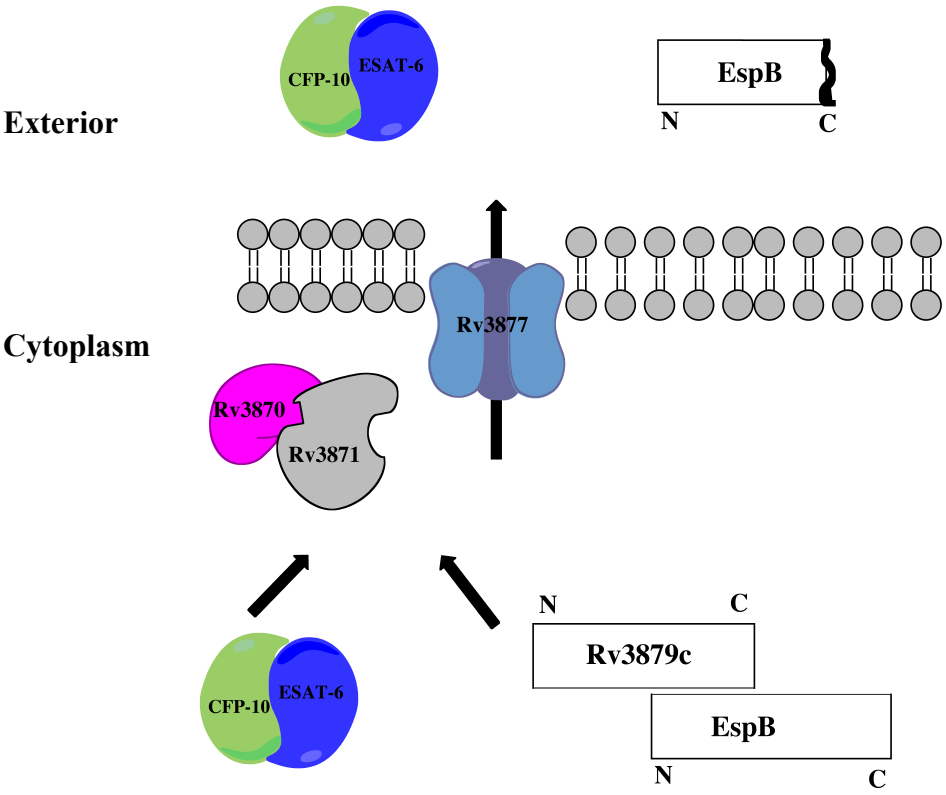


Figure 6.1 EspB secretion model. Core components Rv3870, Rv3871, Rv3879c and Rv3877 and some of the substrates, CFP-10, ESAT-6 and EspB involved in ESX-1 secretion are shown. EspB interacts with Rv3879c and is delivered to the Rv3871 secretion machinery where EspB interacts with ESAT. After secretion, EspB is cleaved on the C-terminus.

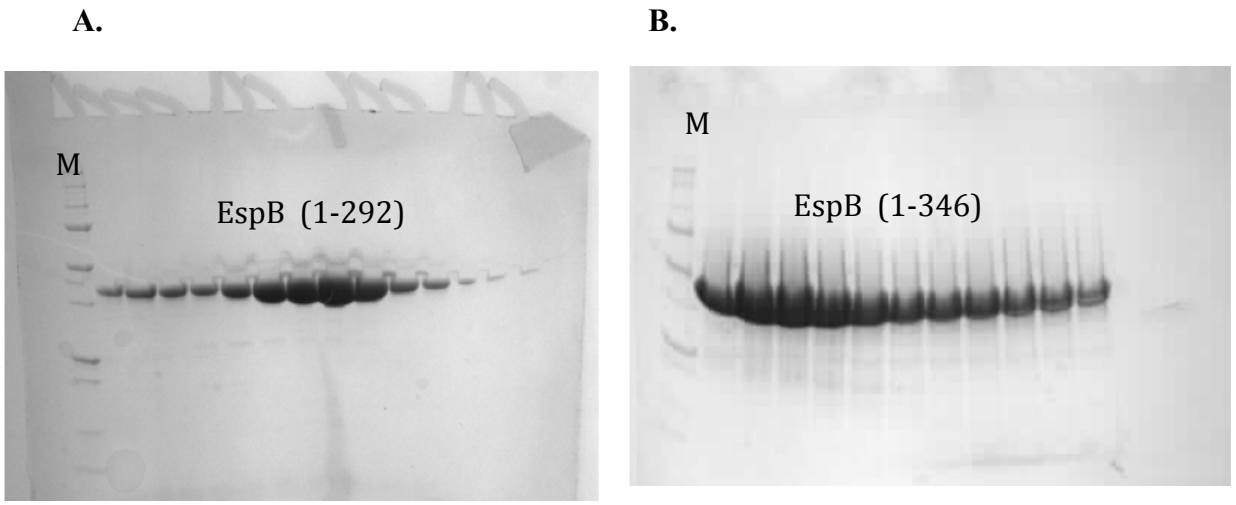


Figure 6.2: Preparative gels after size exclusion chromatography. Lane M in all gels is the protein ladder. **A**. Gel for Se-Met EspB (1-292) with an expected molecular weight of 31.8kDa. **B.** Gel for EspB (1-346) with an expected molecular weight of 36.7 kDa.

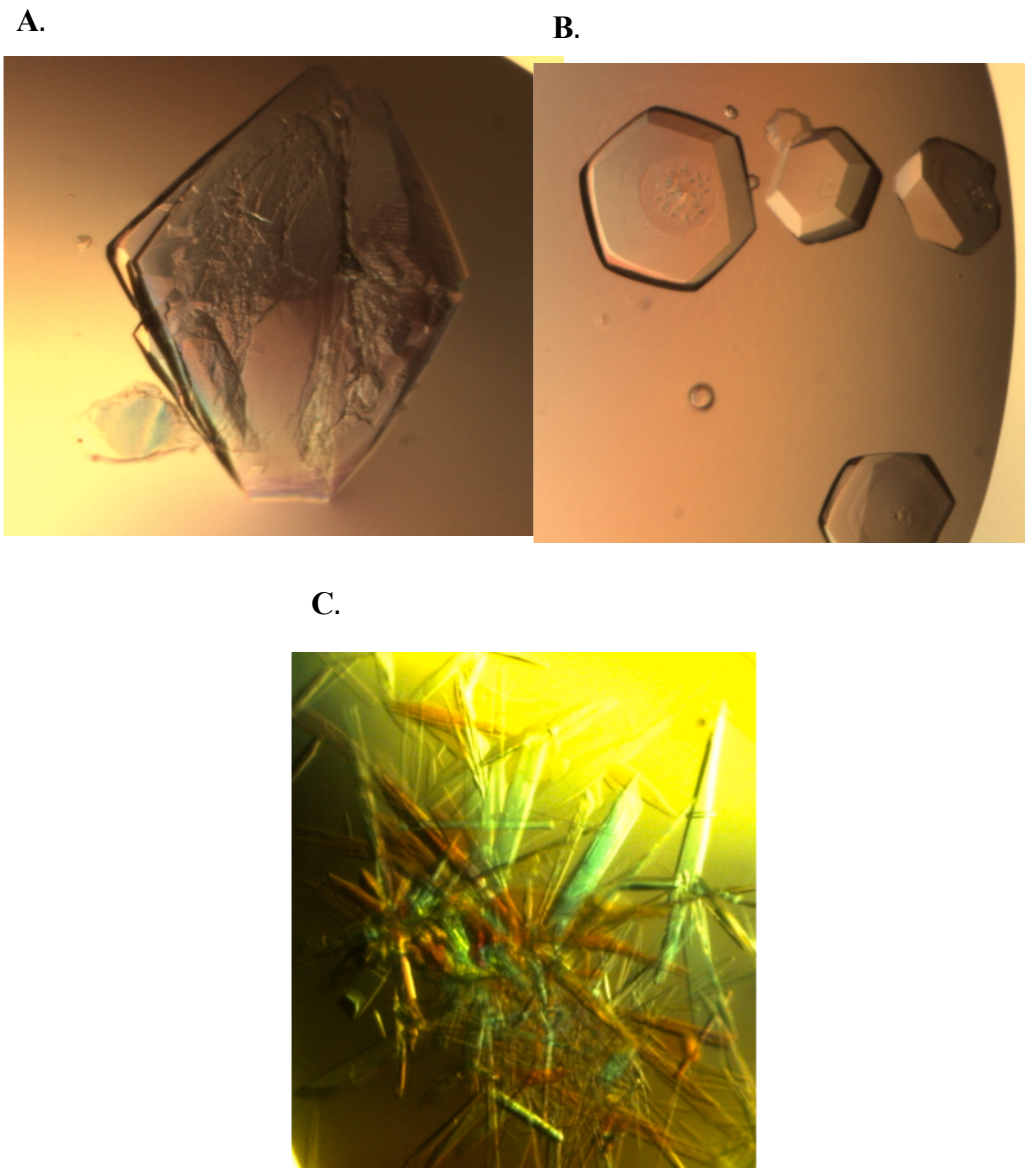


Figure 6.3: Se-Met EspB (1-292) crystals. A. Diamond-shaped crystals at pH 8.5 that diffracted to 2.5Å. **B.** Hexagonal-shaped crystals at pH 6.5 that diffracted to 6.0 Å. **C.** Retangular-shaped crystals at pH 8.5, these crystals did not diffract.

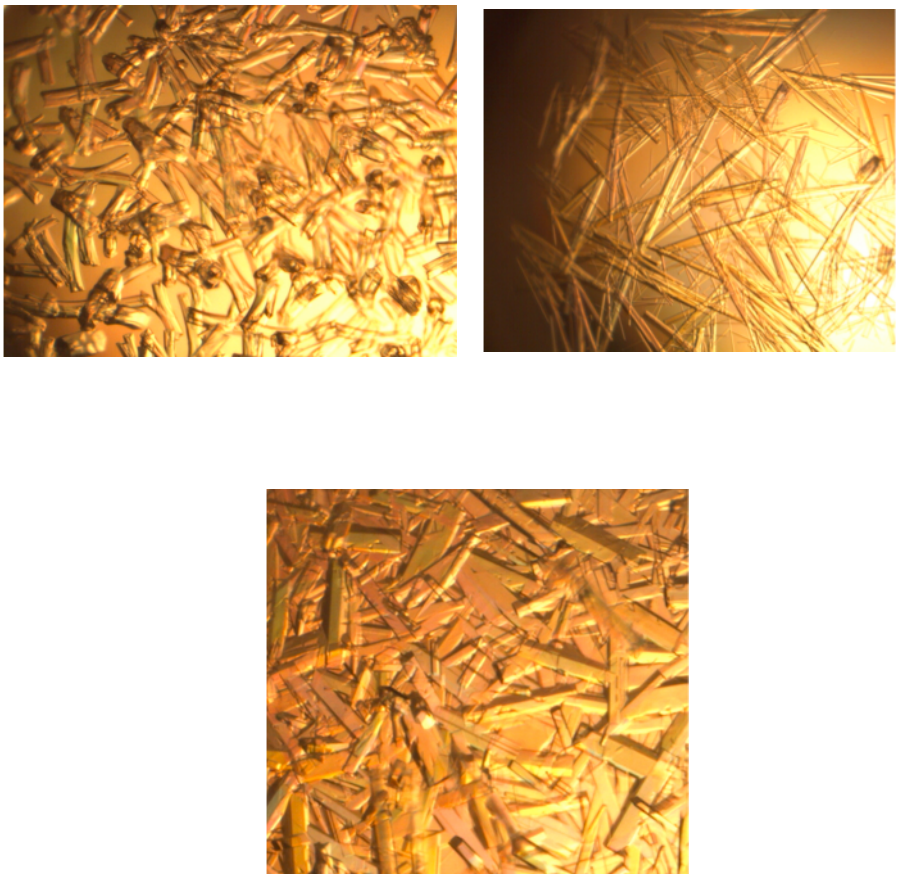


Figure 6.4: EspB (1-346) crystals that diffracted to 6.75 Å.

Table 6.1: Se-Met EspB (1-292) Scalelog file. Diffracted to 2.5Å

Shell Lower limit	Shell Upper limit	Average Redundancy Per Shell
50.00	5.38	3.9
5.38	4.27	3.8
4.27	3.73	3.9
3.73	3.39	3.9
3.39	3.15	3.9
3.15	2.96	3.9
2.96	2.82	3.8
2.82	2.69	3.8
2.69	2.59	3.5
2.59	2.50	3.0
All hkl		3.7

Shell Lower limit	Shell Upper limit	I/Sigma in resolution shells: % of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
50.00	5.38	0.2	0.9	2.0	3.2	5.9	16.0	99.9	0.0	99.9
5.38	4.27	0.6	1.7	3.4	5.2	9.3	24.2	100.0	0.0	100.0
4.27	3.73	1.3	3.2	5.7	8.4	14.3	33.4	100.0	0.0	100.0
3.73	3.39	2.2	5.6	10.1	14.8	23.8	48.4	96.8	3.2	100.0
3.39	3.15	3.9	10.0	17.9	25.4	39.5	67.9	98.3	1.7	100.0
3.15	2.96	6.6	16.6	28.3	39.0	56.3	82.8	98.8	1.2	100.0
2.96	2.82	9.1	22.8	38.4	51.0	68.7	90.0	99.3	0.6	100.0
2.82	2.69	12.6	30.7	49.6	63.5	80.3	95.4	99.7	0.2	99.9
2.69	2.59	16.0	39.5	61.1	74.5	88.0	97.5	99.2	0.0	99.2
2.59	2.50	19.2	46.4	67.8	79.4	89.3	94.8	95.6	0.0	95.6
All hkl		7.2	17.8	28.5	36.5	47.6	65.0	98.8	0.7	99.5

Shell limit	Lower Angstrom	Upper	Average I	Average error	stat.	Norm. Chi**2	Linear R-fac	Square R-fac
50.00	5.38		76.9	4.5	2.0	0.985	0.062	0.071
5.38	4.27		59.1	3.8	2.0	1.316	0.092	0.104
4.27	3.73		51.4	3.6	2.2	1.472	0.111	0.118
3.73	3.39		34.9	2.8	2.2	1.405	0.131	0.133
3.39	3.15		21.6	2.4	2.2	1.128	0.172	0.165
3.15	2.96		13.7	2.4	2.3	1.002	0.247	0.237
2.96	2.82		10.4	2.5	2.4	0.947	0.330	0.304
2.82	2.69		7.4	2.6	2.6	0.887	0.452	0.413
2.69	2.59		5.5	2.9	2.9	0.868	0.605	0.559
2.59	2.50		4.4	3.6	3.5	0.850	0.714	0.649
All reflections			28.6	3.1	2.4	1.098	0.142	0.110

Table 6.2: Se-Met EspB (1-292) Data collection and Refinement Statistics

Data collection	EspB (1-292)
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	173.448,81.703,267.467
α, β, γ (°)	90, 108.77, 90
Molecules/Asym. unit	16
Wavelength (Å)	0.97921
Resolution (Å)	50-2.50
R_{sym} (last shell)	0.142 (0.452)
Completeness (%)	99.5 (95.6)
Redundancy	3.7 (3.0)
Refinement:	
Resolution (Å)	43.8-2.50
R_{work}/R_{free} (%)	21.0/24.9
Mean B-factors $(Å)^2$ Residues	49.8
RMSD Bond lengths (Å)	0.009
RMSD Bond angles (°)	1.176
Ramachandran Plot	
Favored	96.3%
Allowed	0.0%

Table 6.3: Se-Met EspB (1-292) Scalelog file. Diffracted to 2.7Å

Shell Lower limit	Shell Upper limit	Average Redundancy Per Shell
50.00	5.81	6.9
5.81	4.62	6.6
4.62	4.03	6.5
4.03	3.66	6.3
3.66	3.40	6.2
3.40	3.20	6.1
3.20	3.04	6.1
3.04	2.91	6.1
2.91	2.80	6.0
2.80	2.70	6.0
All hkl		6.3

Shell Lower limit	Shell Upper limit	I/Sigma in resolution shells: % of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
50.00	5.81	0.4	0.6	0.9	1.5	2.7	7.5	72.2	27.7	99.9
5.81	4.62	1.5	2.1	3.0	4.1	6.5	16.3	100.0	0.0	100.0
4.62	4.03	2.7	3.4	4.4	5.6	8.6	17.3	55.7	44.3	100.0
4.03	3.66	3.5	4.6	6.3	8.4	12.8	25.7	63.6	36.4	100.0
3.66	3.40	4.4	6.5	9.5	12.9	19.6	37.0	70.2	29.7	99.9
3.40	3.20	6.4	10.0	14.7	19.9	30.1	51.1	79.7	20.1	99.9
3.20	3.04	8.5	14.9	22.8	29.8	42.2	65.0	88.4	11.3	99.7
3.04	2.91	10.3	18.5	28.2	36.9	51.4	74.6	93.1	6.4	99.5
2.91	2.80	12.4	22.9	34.6	45.0	59.9	81.1	95.0	4.1	99.1
2.80	2.70	14.9	28.2	42.3	53.9	69.2	87.1	96.4	1.9	98.3
All hkl		6.5	11.1	16.6	21.7	30.1	46.1	81.4	18.2	99.6

Shell limit	Lower Angstrom	Upper Angstrom	Average I	Average error	Norm. stat.	Linear Chi**2	Square R-fac	R-fac
	50.00	5.81	294.5	15.6	3.6	1.736	0.101	0.119
	5.81	4.62	161.9	10.0	3.3	1.559	0.115	0.127
	4.62	4.03	180.1	8.8	4.1	2.371	0.125	0.134
	4.03	3.66	128.1	6.7	4.0	2.469	0.145	0.151
	3.66	3.40	85.2	5.0	3.9	2.433	0.168	0.167
	3.40	3.20	55.4	4.2	3.8	2.086	0.203	0.189
	3.20	3.04	35.7	3.9	3.7	1.771	0.270	0.241
	3.04	2.91	27.3	3.9	3.8	1.576	0.335	0.289
	2.91	2.80	20.8	3.9	3.8	1.462	0.422	0.360
	2.80	2.70	14.9	3.9	3.8	1.292	0.548	0.463
All reflections			101.5	6.6	3.8	1.880	0.146	0.131

Table 6.4: Hexagonal EspB (1-292) crystals Scalelog file. Diffracted to 6.0 Å

Shell Lower limit	Shell Upper limit	Average Redundancy Per Shell
50.00	12.86	4.3
12.86	10.24	4.7
10.24	8.95	4.7
8.95	8.14	4.5
8.14	7.56	4.5
7.56	7.11	4.6
7.11	6.76	4.6
6.76	6.46	4.6
6.46	6.21	4.7
6.21	6.00	4.7
All hkl		4.6

Shell Lower limit	Shell Upper limit	I/Sigma in resolution shells: % of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
50.00	12.86	2.5	2.9	2.9	3.2	4.6	44.4	97.6	0.0	97.6
12.86	10.24	0.8	1.1	1.5	2.3	4.1	29.6	100.0	0.0	100.0
10.24	8.95	0.7	1.6	2.8	4.1	7.2	33.7	99.7	0.0	99.7
8.95	8.14	1.5	3.9	6.7	10.1	20.0	51.8	100.0	0.0	100.0
8.14	7.56	1.5	5.9	12.6	18.9	32.2	57.0	99.0	1.0	100.0
7.56	7.11	2.8	8.2	16.3	25.0	45.1	73.1	99.2	0.7	99.9
7.11	6.76	5.8	13.8	28.7	40.4	59.9	87.0	99.9	0.1	100.0
6.76	6.46	7.1	19.6	35.3	49.6	68.4	94.1	100.0	0.0	100.0
6.46	6.21	7.1	21.3	38.9	55.1	75.8	95.0	99.7	0.3	100.0
6.21	6.00	9.7	27.0	49.2	63.6	81.4	96.5	100.0	0.0	100.0
All hkl		4.0	10.5	19.5	27.2	39.9	66.2	99.5	0.2	99.7

Shell limit	Lower Angstrom	Upper	Average I	Average error	Norm. stat.	Linear Chi**2	Square R-fac	R-fac
50.00	12.86		99.6	10.6	2.7	1.823	0.128	0.138
12.86	10.24		83.9	7.1	1.7	1.763	0.130	0.144
10.24	8.95		63.2	5.4	1.7	1.958	0.148	0.151
8.95	8.14		32.4	2.9	1.4	2.116	0.166	0.161
8.14	7.56		19.7	1.9	1.4	1.977	0.200	0.220
7.56	7.11		13.1	1.6	1.4	1.726	0.257	0.264
7.11	6.76		8.1	1.5	1.4	1.304	0.345	0.338
6.76	6.46		6.3	1.5	1.5	1.312	0.436	0.431
6.46	6.21		5.7	1.5	1.5	1.040	0.455	0.417
6.21	6.00		4.6	1.6	1.6	1.146	0.601	0.550
All reflections			33.5	3.5	1.6	1.611	0.170	0.151

Table 6.5: EspB (1-346) Scalelog file

Shell Lower limit	Shell Upper limit	Average Redundancy Per Shell
40.00	14.34	3.6
14.34	11.47	3.9
11.47	10.05	4.0
10.05	9.14	4.0
9.14	8.49	4.1
8.49	7.99	4.1
7.99	7.60	4.1
7.60	7.27	4.0
7.27	6.99	4.0
6.99	6.75	3.8
All hkl		4.0

Shell Lower limit	Shell Upper limit	I/Sigma in resolution shells: % of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
40.00	14.34	5.8	13.6	20.3	26.4	33.2	42.4	76.6	23.1	99.7
14.34	11.47	5.5	17.2	28.5	33.6	46.7	62.4	81.4	18.6	100.0
11.47	10.05	10.3	22.5	32.1	39.1	49.1	65.3	88.2	11.4	99.6
10.05	9.14	11.8	22.9	41.2	49.6	59.9	76.3	89.7	10.3	100.0
9.14	8.49	18.8	37.5	53.9	60.2	69.5	87.1	96.1	3.9	100.0
8.49	7.99	13.8	38.7	53.3	64.4	75.5	88.5	95.8	4.2	100.0
7.99	7.60	13.5	39.4	58.2	67.3	80.5	90.8	97.6	2.4	100.0
7.60	7.27	15.6	41.6	57.2	70.4	79.2	92.4	99.6	0.0	99.6
7.27	6.99	24.2	42.7	60.8	69.6	79.6	90.0	96.5	2.3	98.8
6.99	6.75	15.7	41.0	58.6	67.9	80.3	94.8	96.8	0.0	96.8
All hkl		13.3	31.2	45.8	54.1	64.6	78.2	91.5	8.0	99.5

Shell limit	Lower Angstrom	Upper	Average I	Average error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac
40.00	14.34		19.3	1.0	0.7	1.159	0.057	0.033
14.34	11.47		11.3	0.6	0.6	1.018	0.081	0.066
11.47	10.05		8.7	0.6	0.5	0.974	0.089	0.055
10.05	9.14		6.3	0.6	0.5	0.958	0.134	0.085
9.14	8.49		2.8	0.5	0.5	0.996	0.213	0.143
8.49	7.99		3.1	0.5	0.5	0.948	0.251	0.158
7.99	7.60		2.4	0.5	0.5	0.709	0.263	0.172
7.60	7.27		1.6	0.6	0.5	0.855	0.423	0.294
7.27	6.99		1.8	0.6	0.6	0.961	0.332	0.204
6.99	6.75		1.2	0.6	0.6	0.677	0.463	0.402
All reflections			6.1	0.6	0.6	0.929	0.136	0.057

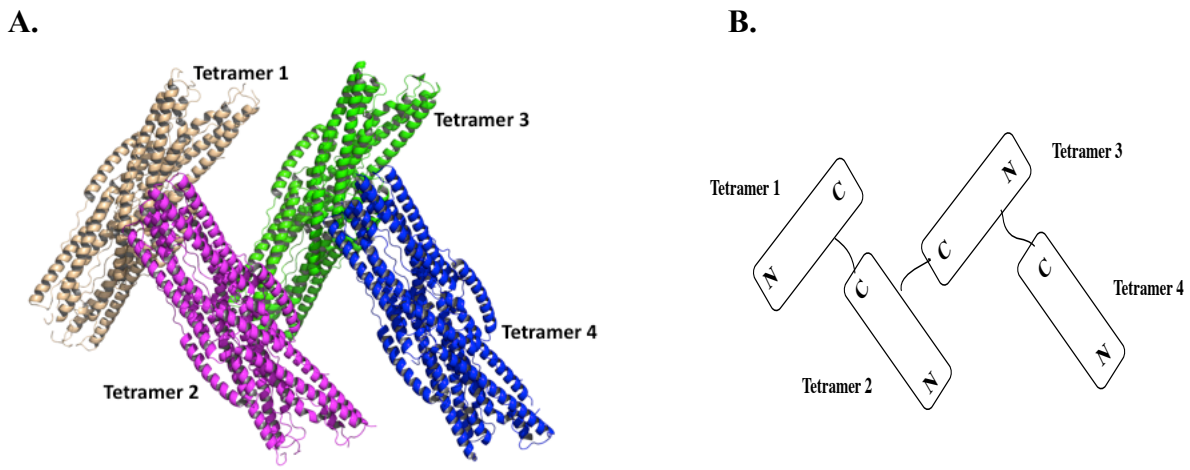


Figure 6.5: Overall structure of EspB (1-292). Tetramer of tetramers, two tetramers: 1 and 3; and 2 and 4 are parallel and adjacent tetramers interact via the C-terminal of one tetramer and the C-lobe of another tetramer.

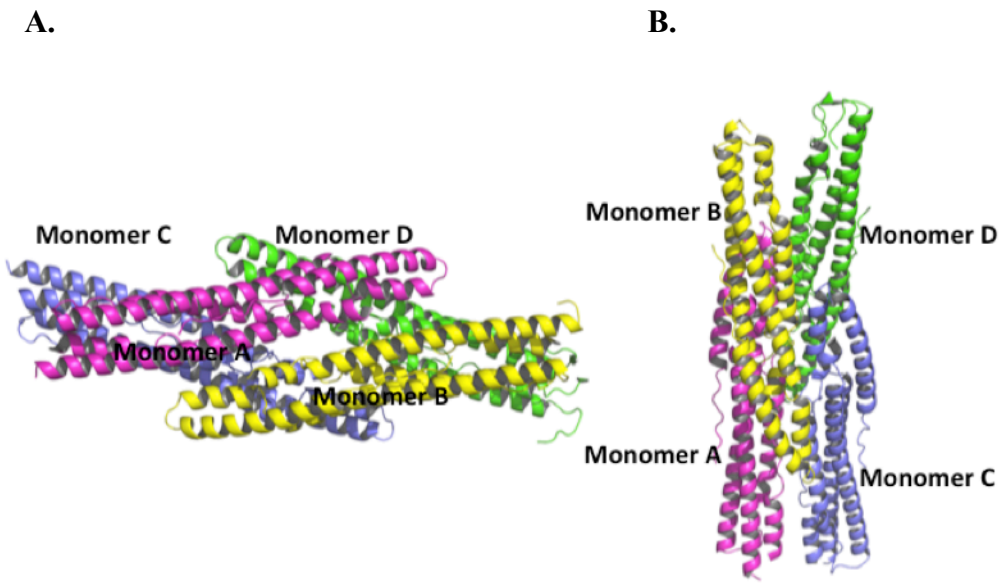


Figure 6.6: Overall structure of one homotetramer. Monomer A is in magenta, monomer B in yellow, monomer C in blue and monomer D in green. Monomers A and B are anti-parallel and so are monomers C and D. Monomers A and D as well as monomers B and C interact via the C-lobe helices.

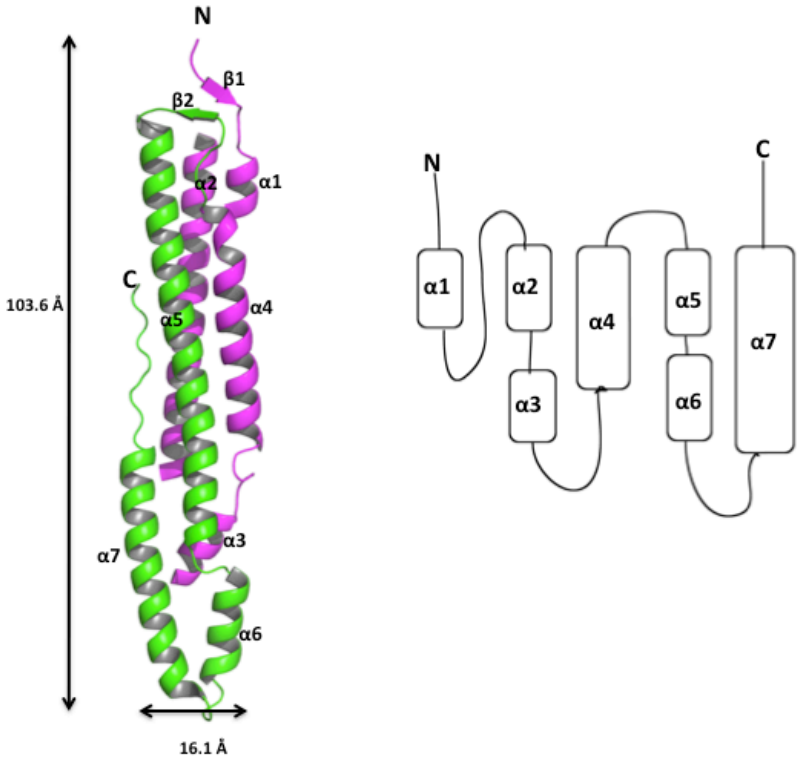


Figure 6.7: **Overall monomeric structure**. The monomer is mostly helical and cylindrical in overall shape.

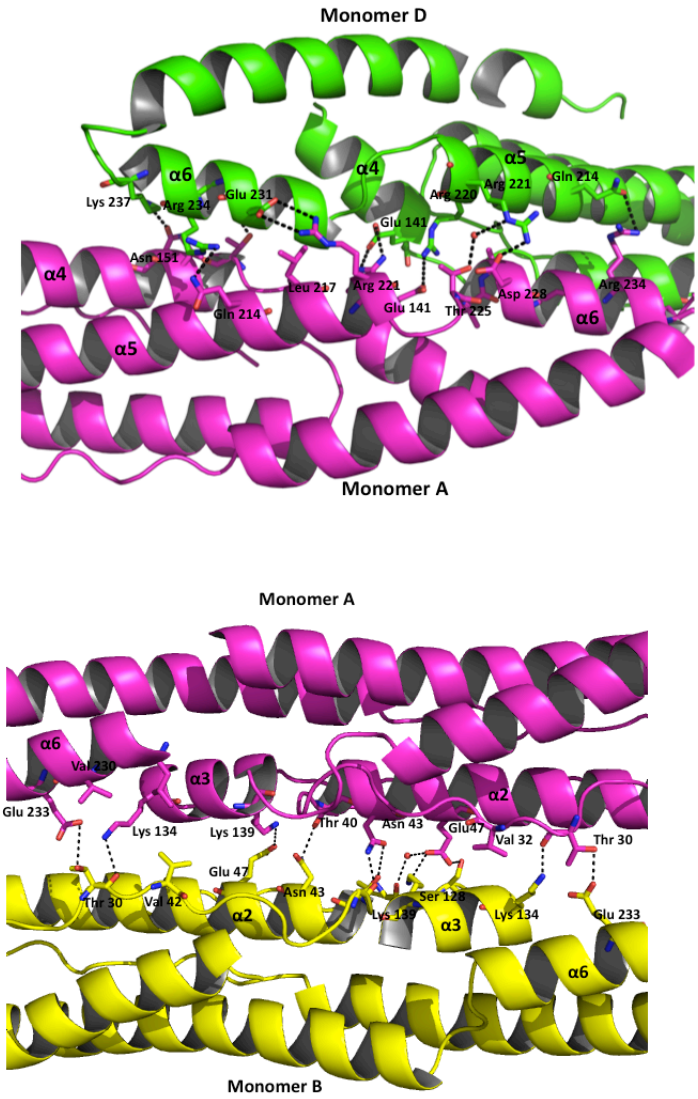


Figure 6.8: **Interacting interfaces**. Black dashed lines are hydrogen bonds and red spheres are waters. **A.** Interface between the C-lobe helices (α4, α5 and α6). **B.** Interface between the anti-parallel monomers involving α2, α3, and α6.

Chapter 7: Summary and Conclusion

7.1 Summary

Regulating the rate of ribosome biogenesis is an accessible avenue to modulating cell proliferation, especially in abnormal proliferative cells such as cancer cells. One approach of achieving regulation of cell proliferation is by targeting proteins involved in ribosome biogenesis. This approach has precedence in studies that show deletion of a ribosomal protein, S6 in mice results in decreased cell proliferation yet cell growth is sustained in response to nutrient stimuli. The work described here takes the same approach by targeting Rio1 kinase that is required for both maturation of 18S rRNA during ribosome biogenesis and cell cycle progression. Towards this approach, potential inhibitors were tested for binding to *Archaeoglobus fulgidus* Rio1 (afRio1).

Small molecules were selected based on a previously solved crystal structure of afRio1 bound to ATP/Mn^{2+} that revealed an accessible pocket near the N7 position on the adenine moiety of ATP. Toyocamycin, sangivamycin, 7-methylinosine and ATP/Mg^{2+} were tested for binding to afRio1 using a thermofluor assay. Toyocamycin, an adenosine analog with a carbon and a cyano group at the N7 position demonstrated the highest affinity for afRio1. Despite the lack of a triphosphate moiety, toyocamycin stabilized afRio1 and bound to afRio1 with a higher affinity than the native ligand, ATP/Mg^{2+}. Co-crystallization of afRio1 with toyocamycin resulted in a crystal structure of afRio1-toyocamycin that showed occupancy of toyocamycin in the ATP binding pocket. Interactions between toyocamycin and afRio1 revealed the metal binding residue, Asp 212 that interacts

the γ-phosphate of ATP via a metal-mediated interaction, was able to hydrogen bond with the 5'OH of toyocamyin despite the lack of a triphosphate moiety. This suggests that the triphosphate moiety may not be critical for tight binding to afRio1, as toyocamycin binds with a higher affinity to afRio1 than ATP. The presence of the cyano group at the N7 position on toyocamycin may also facilitate tighter binding. Also additional hydrogen bonds, both direct and water-mediated, that are observed in afRio1-toyocamycin complex but not the ATP bound structure may compensate for the lack of the triphosphate moiety.

Binding of toyocamycin to the ATP binding pocket necessitated inhibition studies that were carried out to validate competitive inhibition that occurs when the inhibitor binds to the substrate's pocket. Competitive inhibition is characterized by a change in the affinity of the enzyme for the substrate (K_m) and no change in the enzyme's maximum catalytic rate (V_{max}). AfRio1 inhibition studies revealed unexpected results of a change in both the K_m and $V_{max.}$ Though crystallographic studies suggested strict competitive inhibition, steady state kinetic studies suggested that uncompetitive inhibition mechanism was also in play. This required assuming a mechanism that could account for both competitive and uncompetitive inhibition. A mixed inhibition model that is characterized by a change in both K_m and V_{max} was adopted in an effort to explain the results. Mixed inhibition occurs via allosteric modulation where the inhibitor binds to a different site than the substrate's binding pocket or through modulation of binding sites via formation of oligomers. Since co-crystallization studies showed occupancy of toyocamycin in the ATP binding pocket, the only other notion to explore was the formation of oligomers. In addition to

inhibition of afRio1 by toyocamycin, another phenomenon known as substrate inhibition that occurs in the presence of multiple binding sites was observed at high concentrations of ATP. In order to explain both mixed inhibition and substrate inhibition, sedimentation equilibrium experiments were carried out to determine whether oligomers were being formed. In the presence of toyocamycin, a monomer-tetramer with an average molecular weight of 71.8 kDa was observed in solution. In the presence of ATP, a monomer-dimer with an average molecular weight of 53.6 kDa was observed whereas when phosphorylated afRio1 has a molecular weight of 34.1 kDa and fits to a single species model that is monomeric. In the presence of ATP or toyocamycin, the monomeric afRio1 (phosphorylated afRio1) average molecular weight increases to 47.6 kDa and 46.2 kDa respectively and both fit to a monomer-trimer model. Taken together this data suggests that ATP and toyocamycin promote oligomerization and both ligands stabilize a less catalytically active oligomer. Crystal packing analysis of the different oligomers revealed that ATP and toyocamycin stabilized the same type of dimer and have a similar interacting interface, different from the adenosine bound structure. This was somewhat unexpected since toyocamycin is an adenosine analog, and lacks the triphosphate moiety. Despite this, toyocamycin favors the ATP bound bound conformation and the data supported and was consistent with both mixed inhibition and substrate inhibition. This data also links autophosphorylation and oligomerization as a mechanism that afRio1 uses to modulate afRio1 catalytic activity.

There is currently no crystal structure of eukaryotic Rio1 kinase and part of my research was to determine the crystal structure of human Rio1 kinase, RioK1.

Several strategies were employed in cloning, expressing and purifying a construct of RioK1. This included limited proteolysis and surface entropy mutagenesis complemented by mass spectrometry analysis. RioK1 residues 143-494, was the most stable fragment and resulted in the first crystal structure of RioK1 with ATP/Mg^{2+} bound. Analysis of the interaction between RioK1 and ATP/Mg^{2+} revealed similar hydrophobic contacts and hydrogen bonds as those observed between afRio1 and ATP/Mn^{2+}. An alignment of RioK1-ATP/Mg^{2+} with afRio1-ATP/Mn^{2+} had an RMSD of 0.894 Å over 158 residues with shifts observed on the N-lobe's twisted β-sheet, the nucleotide binding loop, the hinge region and the metal binding loop. In the alignment, the γ-phosphate of ATP in the RioK1 (143-494)-ATP /Mg^{2+} is oriented differently compared to the afRio1-ATP/Mn2+ structure. In addition, the distance between the γ-phosphate and the metal binding residue is 3.3 Å in the RioK1-ATP complex whereas in afRio1-ATP/Mn^{2+} this distance is 5.1 Å. The difference in this distance can be attributed to the shift in the metal binding loop and is also suggestive of a difference in the catalytic rate activity between eukaryotic Rio1 and afRio1. Crystal packing analysis using PDBe PISA (protein interfaces, surface and assemblies) server, resulted in two potential biological interfaces, one between the two molecules in the asymmetric unit and another interface between two symmetry related molecules. Mutational studies, oligomeric characterization and functional studies would aid in determining the biological relevant interface.

In addition to structure determination of RioK1, there was interest in determining whether biochemical characteristics determined for afRio1 were conserved in RioK1. RioK1 oligomeric state characterization, affinity for potential

Rio1 inhibitors and substrate inhibition by high concentrations of ATP were analyzed. In all three cases, RioK1 showed a similar trend as afRio1, in the presence of ATP and adenosine, RioK1 formed a dimer and tetramer respectively and when phosphorylated RioK1 was a mixture of monomer-dimer in solution suggesting that RioK1 may employ the same inhibition mechanism as afRio1 and phosphorylation may regulate catalytic activity. At high concentrations of ATP above 9.0 μM, RioK1 phosphorylation activity was inhibited, as was the case with afRio1 suggesting that both proteins may employ the same substrate inhibition mechanism through stabilization of a less catalytically active form. Finally during analysis of RioK1's affinity for potential inhibitors, toyocamycin bound to RioK1 with a higher affinity than ATP/Mg^{2+}. This is similar to what was observed with afRio1 suggesting that toyocamycin may bind and interact with RioK1 as it does with afRio1. Interestingly both eukaryotic and archaeal Rio1 kinase demonstrate conservation in both structural features and biochemical characteristics despite over a billion years of evolution.

In addition to studying Rio1 kinase, this thesis presents the first crystal structure of a mycobacterium tuberculosis virulent factor known as EspB. The EspB (1-292) structure shows a conserved arrangement of tetramers at different pH conditions and in different crystal forms. The crystal structure of a larger fragment, EspB (1-346), also shows the same arrangement of tetramers suggesting that the arrangement has a possible biological significance.

7.2 Future Work and Direction

Co-crystallization studies of afRio1 with toyocamycin provide a first generation class of inhibitors that will aid in designing more Rio1-specific inhibitors. This would include: a) exploring different substitution groups at the N7 position as well as the base b) exploring substitution on the ribose, for instance on the 3'OH that has the closest residue 6.2Å away, c) functionalizing the triphosphate moiety by introducing non-hydroylzable groups (NH_2, S, CH) between the β and γ phosphates or d) molecules that mimic the transition state. Identifying Rio1 substrates is also key in determining Rio1-specific peptide inhibitors, a tool that has been used to determine kinase inhibitors. An example of a kinase peptide inhibitor is Kemptide that inhibits Protein Kinase A (PKA). These studies would also require characterizing the mode of inhibition.

A co-crystal structure of the human Rio1 (RioK1) with toyocamycin bound would aid in determining whether toyocamycin maintains similar interactions with RioK1 as with afRio1. In addition to crystallization, inhibition studies would be essential to determine both afRio1 and RioK1 share the same inhibition mechanism. This would be beneficial in designing Rio1-specific inhibitors since those that inhibit afRio1 would possibly inhibit RioK1. This would also serve as a valuable tool in identifying Rio1 substrates by using radiolabeled toyocamycin triphosphate or by identifying proteins with diminished phosphorylation activity in he cells in the presence of toyocamycin. Both methods have been explained explicitly in chapter 2. Recent studies that have identified an interaction of RioK1 with protein arginine

methyltransferase 5 and nucleolin provide the opportunity to study the crystal structures of the complexes and better characterize the interaction.

In addition to Rio1 studies, work done with EspB requires determination of EspB oligomeric states in the presence of different conditions that mimic phagosomal environment, the crystal structures of proteolytic fragments of EspB as a result of cleavage of Myc1 after secretion from the cytoplasm as well as determining the role of these fragments. The crystal structure of EspB bound to ESAT-6, the ATPase Rv3871 and Myc1 protease would be also be essential in better understanding the ESX-1 secretion system.

REFERENCES:

1. Koyama, H. I., M.; Nishizawa, Y.; Ohno, S.; and Moril, H.; . (1994) Protein kinase C is involved in 24-hydroxylase gene expression induced by 1,25(OH)2D3 in rat intestinal epithelial cells. *J. Cell Biochem. 55*, 230-240.
2. Bouchoux, C. H., G.; Grenetier, S.; Carles, C.; Riva, M.; and Gogue, V.;. (2004) CTD kinase I is involved in RNA polymerase I transcription, *Nucleic Acid Research 32*, 5851-5860.
3. Kawasaki, L. S., O.; Shiozaki,K.; and Aguire,J. (2002) SakA MAP kinase is involved in stress signal transduction, sexual development and spore viability in Aspergillus nidulans, *Mol. Microbiol. 45*, 1153-1163.
4. Cappellini, A. T., G.; Zweyer,M.; Bortul,R.;Tazzari,P.L.; Billi,A.M.; Fala,F.; Cocco,L.; and Martelli,A.M. (2003) The phosphoinositide 3-kinase/Akt pathway regulates cell cycle progression of HL60 human leukemia cells through cytoplasmic relocalization of the cyclin-dependent kinase inhibitor p27(Kip1) and control of cyclin D1 expression., *Leukemia 17*, 2157-2167.
5. Heindl, K. a. M., J. (2010) Nol9 is a novel polynucleotide 5'-kinase involved in ribosomal RNA processing. *The EMBO Journal 24*, 4161-4171.
6. al, H. S. Y. a. e. (1997) P-TEFb kinase is required for HIV Tat transcriptional activation in vivo and in vitro, *Gene & Development 11*, 2633-2644.
7. Sophie, L. a. e. a. (2001) Indirubins Inhibit Glycogen Synthase Kinase-3b and CDK5/P25, Two Protein Kinases Involved in Abnormal Tau Phosphorylation in Alzheimer's Disease, *J. Biol. Chem. 276*, 251-260.
8. Berger, H. A. T., M.S.; and Welsh,M.J. (1993) Regulation of the Cystic Fibrosis Transmembrane Conductance Regulator C1- Channel by Specific Protein Kinases and Protein Phosphatases, *J. Biol. Chem. 268*, 2037-2047.
9. Yadav, V. a. D. M. F. (2011) Fyn is induced by Ras/PI3K/Akt signaling and is required for enhanced invasion/migration, *Mol. Carcinog. 50*, 346-352.
10. Hanks, S. K., and Hunter, T. (1995) The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification, *FASEB Journal 9*, 576-596.
11. Cleland, W. W. a. H., A.C. (1995) Mechanisms of Phosphoryl and Acyl Transfer, *FASEB Journal 9*, 1585-1594.
12. Madhusudan, E. A. T., N. H. Xuong, J. A. Adams, L. F. Ten Eyck, S. S. Taylor, and J. M. Sowadski. (1994) cAMP-dependent protein kinase: crystallographic insights into substrate recognition and phosphotransfer.*Protein Sci 2*, 176-187.
13. Zheng, J., Knighton,D.R., Eyck, L., Karlsson,R., Xuong,N.H., Taylor, S.S. and Sowadski,J.M., . (1992) Crystal Structure of the Catalytic Subunit of cAMP-dependent Protein Kinase Complexed with MgATP and Peptide Inhibitor., *Biochemistry 32*, 2154-2161
14. Montenegro, M., Garcia-Viloca, M.,Gonzalez-Lafont,A., Lluch,M.J., . (2007) Comparative study of the prereactive protein kinase A Michaelis complex with Kemptide substrate, *J.Comput Aided Mol Des 21*, 603-615.
15. Montenegro, M., Garcia-Viloca, M.,Gonzalez-Lafont,A., Lluch,M.J., . (2011) A QM/MM study of the phosphoryl transfer to the Kemptide

substratecatalyzed by protein kinase A. The effect of the phosphorylation state of the protein on the mechanism, *Physical Chemistry Chemical Physics 13*, 530-539.

16. Manning, G., Whyte,D.B., Martinez,R., Hunter,T., S. Sudarsanam. (2002) The Protein Kinase Complement of the Human Genome, *Science 298*, 1912-1934.
17. Kostich, M., English,J., Madison,V., Gheyhas,F.,Wang, L., Qiu,P., Greene,J. and Laz, T.M. (2002) Human members of the eukaryotic protein kinase family, *Genome Biology 3*, 1-12.
18. LaRonde-LeBlanc, N., and Wlodawer, A. (2005) The RIO kinases: An atypical protein kinase family required for ribosome biogenesis and cell cycle progression, *Biochimica et Biophysica Acta 1754*, 14-24.
19. Manning, G., Plowman,G.D., Hunter,T., and Sudarsanam,S. (2002) Evolution of protein kinase signaling from yeast to man, *TRENDS in Biochemical Sciences 27*, 514-520.
20. Leonard, C. J., Aravind,L., and Koonin,E. . (1998) Novel Families of Putative Protein Kinases in Bacteria and Archaea: Evolution of the ''Eukaryotic'' Protein Kinase Superfamily, *Genome Research 8*, 1038–1047.
21. Goffeau, A., Barrell,B. G. , Bussey, H. ., Davis,R. W., Dujon, B.,Feldmann, H., Galibert, F., Hoheisel,J. D., Jacq, C. , Johnston, M.,Louis, E. J. , Mewes, H. W.,Murakami,Y. , Philippsen,P., Tettelin,H. , and Oliver,S. G. . (1996) Life with 6000 Genes, *Science 274*, 546-567.
22. Angermayr, M., and Bandlow, W. (1997) The Type of Basal Promoter Determines the Regulated or Constitutive Mode of Transcription in the Common Control Region of the Yeast Gene Pair GCY1/RIO1, *J.Biol.Chem 272*, 31630-31635.
23. Angermayr, M., Schwerdtfeger,K.,Bandlow, W. (2003) A Nucleosome-Free dG-dC-Rich Sequence Element Promotes Constitutive Transcription of the Essential Yeast RIO1 Gene, *Biol.Chem. 384*, 1287-1292.
24. LaRonde-LeBlanc, N., Guszczynski,T., Copeland, T., and Wlodawer,A. (2005) Structure and activity of the atypical serine kinase Rio1, *FEBS Journal 272*, 3698-3713.
25. LaRonde-LeBlanc, N., Guszczynski,T., Copeland, T., and Wlodawer,A. (2005) Autophosphorylation of Archaeoglobus fulgidus Rio2 and crystal structures of its nucleotide–metal ion complexes, *FEBS Journal 272*, 2800-2810.
26. LaRonde-LeBlanc, N., and Wlodawer, A. (2004) Crystal Structure of A. fulgidus Rio2 Defines a New Family of Serine Protein Kinases, *Structure 12*, 1585-1594.
27. LaRonde-LeBlanc, N., and Wlodawer, A. (2005) A Family Portrait of the RIO Kinases, *J.Biol.Chem 280*, 37297–37300.
28. Srinivasan, N., and Krupa, A. (2002) Lipopolysaccharide phosphorylating enzymes encoded in the genomes of Gram-negative bacteria are related to the eukaryotic protein kinases, *Protein Science 11*, 1580-1584.
29. Hanks, S. K., and Hunter, T. (1988) The Protein Kinase Family: Conserved Features and Deduced Phylogeny of the Catalytic Domains, *Science 241*, 42-52.

30. Zheng, J., Trafny, E.A., Knighton, D.R., Xuong,N.,Taylor,S.S., Ten Eyck, L.F., and Sowadski,J.M. (1993) 2.2 Å refined crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MnATP and a peptide inhibitor, *Acta Crystallographica Section D 49*, 362-265.
31. Angermayr, M., and Bandlow,W. (2002) RIO1, an extraordinary novel protein kinase, *FEBS Letters 524*, 31-36.
32. Angermayr, M. R., A. and Bandlow, W. (2002) Yeast Rio1p is the founding member of a novel subfamily of protein serine kinases involved in the control of cell cycle progression, *Molecular Microbiology 44*, 309-324.
33. Vanrobays, E., Gelugne,J-P., Gleizes, P-E., and Caizergues-Ferrer,M. . (2003) Late Cytoplasmic Maturation of the Small Ribosomal Subunit Requires RIO Proteins in Saccharomyces cerevisiae, *Molecular and Cellular Biology 23*, 2083-2095.
34. Vanrobays, E., Gelugne,J-P., Gleizes, P-E., Bousquet-Antonelli, C., Noaillac-Depeyre,J., Caizergues-Ferrer,M. and Gelugne, J-P. (2001) Processing of 20S pre-rRNA to 18S ribosomal RNA in yeast requires Rrp10p, an essential non-ribosomal cytoplasmic protein, *EMBO Journal 20*, 4204-4213.
35. Soudet, J., Gelugne, J-P., Belhabich-Baumas, K., Caizergues-Ferrer,M., and Mougin, A. . (2010) Immature small ribosomal subunits can engage in translation initiation in Saccharomyces cerevisiae, *EMBO Journal 29*, 80-92.
36. Schluenzen, F., Tocilj,A., Zarivach,R., Harms,J., Gluehmann,M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., and Yonath,A.,. (2000) Structure of Functionally Activated Small Ribosomal Subunit at 3.3 A° Resolution, *Cell 102*, 615-623.
37. Ban, N., Nissen,P., Hansen,J., Moore, P.B., and Steitz, T. (2000) The Complete Atomic Structure of the Large Ribosomal Subunit at 2.4 Å Resolution, *Science 289*, 905-919.
38. Dunkle, J. A., Xiong,L., Mankin, A.S., Cate, J.H.D. (2010) Structures of the Escherichia coli ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action, *PNAS 107*, 17152-17157.
39. Ben-Sham, A., Jenner, L., Yusupova, G., and Yusupov, M. (2010) Crystal Structure of the Eukaryotic Ribosome, *Science 330*, 1203-1208.
40. Taylor, D. J., Devkota,B., Huang,A.D., Topf,M., Narayanan,E., Sali,A., Harvey,S.J., and Frank J. (2009) Comprehensive Molecular Structure of the Eukaryotic Ribosome, *Structure 17*, 1591-1604.
41. Schuwirth, B. S., Borovinskaya,M.A., Hau,C.W., Zhang, W., Vila-Sanjurjo,A., Holton,J.M., Cate, J.H.D. (2005) Structures of the Bacterial Ribosome at 3.5 A° Resolution, *Science 310*, 827-833.
42. Zhang, W., Dunkle, J.A., Cate, J.H.D. (2009) Structures of the Ribosome in Intermediate States of Ratcheting, *Science 325*, 1014-1017.
43. Jenner, L. B., Demeshkina,N.,Yusupov,G., and Yusupov, M. (2010) Structural aspects of messenger RNA reading frame maintenance by the ribosome, *Nature Structural and Molecular Biology 17*, 555-560.
44. Venema, J. a. T., D. . (1999) Ribosome Synthesis in Saccharomyces Cerevisiae, *Annu. Rev. Genet. 33*, 261-311.

45. Fromont-Racine, M., Senger, B., Saveanu, C., and Fasiolo, F. (2003) Ribosome assembly in eukaryotes, *Gene 313*, 17-42.
46. Schaefer, T., Straub, D., Petfalski,E., Tollervey, D., and Hurt, E. . (2003) The path from nucleolar 90S to cytoplasmic 40S pre-ribosomes, *EMBO Journal 22*, 1370-1380
47. Guderian, G., Peter,C., Wiesner, J., Sickmann, A., Schuleze-Osthoff, K., Fischer,U., and Grimmler, M. (2011) RioK1, a new interactor of protein arginine methyltransferase 5 (PRMT5), competes with pICln for binding and modulates PRMT5 complex composition and substrate specificity, *J.Biol.Chem. 286*, 1976-1986.
48. Glnlsty, H., Slcard, H., Roger, B., and Bouvet, P. (1999) Structure and functions of nucleolin, *Journal of Cell Science 112*, 761-772.
49. Derenzini, M., Trere, D., Pession, A., Montanaro, L., Sirri, V., and Ochs, R. L. (1998) Nucleolar function and size in cancer cells, *American Journal of Pathology 152*, 1291-1297.
50. Moss, T. (2004) At the crossroads of growth control; making ribosomal RNA, *Current Opinion in Genetics & Development 14*, 210-217.
51. Derenzini, M., Trere, D., Pession, A., Govoni, M., Sirri, V., and Chieco, P. (2000) Nucleolar size indicates the rapidity of cell proliferation in cancer tissues, *Journal of Pathology 191*, 181-186.
52. Volarevic, S., Stewart, M.J., Ledermann, B., Zilberman,F., Terracciano, L., Montini, E., Grompe,M., Kozma, S.C., and Thomas, G. (2000) Proliferation, But Not Growth, Blocked by Conditional Deletion of 40S Ribosomal Protein S6, *Science 208*, 2045-2047.
53. Line, A., Slucka, Z., Strengrevics, A., Silina, K., and Rees, R.C. (2002) Characterisation of tumour-associated antigens in colon cancer, *Cancer Immunology and Immunotherapy 51*, 574-582.
54. Kendrew, J. C. B., G. Dintzis, H.M., Parrish, R.G., Wyckoff, H., and Phillips, D.C. . (1958) A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis, *Nature 4610*, 662-666.
55. Sheldrick, G. M. (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification, *Acta Crystallographica Section D 66*, 479-485.
56. Dodson, E. (2003) Is it jolly SAD, *Acta Crystallographica Section D 59*, 1958-1965.
57. Wang, J. W., Chen, J.R., Gu, Y.X., Zheng, C.D. anf Fan, H.F. . (2004) Direct-method SAD phasing with partial-structure iteration: towards automation, *Acta Crystallographica 60*, 1991-1996.
58. Carver, T. E., and et al. (2005) Decrypting the Biochemical Function of an Essential Gene from Streptococcus pneumoniae Using ThermoFluor® Technology, *Journal of Biological Chemistry 280*, 11704-11712.
59. Nishiruma, H., Katagiri, K., Sato,K., Hayama,M., and Shimaoka, N.,. (1956) Toyocamycin, a new anti-candida antibiotics, *Journal of Antiobiotics 9*, 60-62.

60. Tavitan, A., Uretsky, S.C., and George, A.C.S. . (1968) Selective Inhibition of Ribosomal RNA Synthesis in Mammalian Cells, *Biochim. Biophys. Acta 157*, 33-42.
61. Renau, T. E., Wotring, L.L., Drach, J.C., and Townsend, L.B. (1996) Synthesis of Non-nucleoside Analogs of Toyocamycin, Sangivamycin, and Thiosangivamycin: Influence of Various 7-Substituents on Antiviral Activity, *Journal of Medicinal Chemistry 39*, 873-880.
62. Tavitian, A., Vretsku S.C. and George, A.C.S. . (1969) The Effect of Toyocamycin on Cellular RNA Synthesis, *Biochim. Biophys. Acta 179*, 50-57.
63. Heine, U. (1969) Electron Microscopic Studies on HeLa Cells Exposed to the Antibiotic Toyocamycin, *Cancer Research 29*, 1875-1880.
64. Riman, J., Sverak,L., Langlois, A.J., Bonar, R.A., and Beard, J.W. (1969) Influence of Toyocamycin on RNA Synthesis in Chick Embryo Cells Noninfected and Infected with Strain MC29 Avian Leukosis Virus', *Cancer Research 29*.
65. Cohen, M. B., and Glazer, R.I. (1985) Comparison of the Cellular and RNA-Dependent Effects of Sangivamycin and Toyocamycin in Human Colon Carcinoma Cells, *Molecular Pharmacology 27*, 349-355.
66. lapalucci-Espinoza, S., Cereghini,S., and Franze-Fernindez,T. (1977) Regulation of Ribosomal RNA Synthesis in Mammalian Cells: Effect of Toyocamycint, *Biochemistry 16*, 2885-2889.
67. Nishioka, H., Sawa,T., Hamada,M., Shimura, N., lmoto,M., and Umezawa,K. (1990) Inhibition of Phosphatidylinositol Kinase by Toyocamycin, *The Journal of Antibiotics 12*, 1586-1589.
68. Caballero, J., Fernandez,M., and Gonzalez-Nilo, F.D. (2008) A CoMSIA study on the adenosine kinase inhibition of pyrrolo[2,3-d]pyrimidine nucleoside analogues, *Bioorganic and Medicinal Chemistry 16*, 5103-5108.
69. Long, M. C., and Parker, W.B. . (2006) Structure–activity relationship for nucleoside analogs as inhibitors or substrates of adenosine kinase from Mycobacterium tuberculosis I. Modifications to the adenine moiety, *Biochemical Pharmocology 71*, 1671-1682.
70. Palczewski, K., Kahn,N., and Hargrave,P. A. (1990) Nucleoside Inhibitors of Rhodopsin Kinase, *Biochemistry 29*, 6276-6282.
71. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray Diffraction Data Collected in Oscillation Mode, *Methods in Enzymology 276*, 307-326.
72. Project, C. C. (1994) The CCP4 Suite: Programs for Protein Crystallography, *Acta Cryst. D50*, 760-763.
73. Vagin, A., Teplyakov,A. (1997) MOLREP: an automated program for molecular replacement, *Journal of Applied Crystallography 30*, 1022-1025.
74. Vagin, A. A., Steiner, R.S., Lebedev, A.A., Potterton, L., McNicholas, S., Long, F. and Murshudov, G.N. (2004) REFMAC5 dictionary: organisation of prior chemical knowledge and guidelines for its use, *Acta Cryst. D60*, 2284-2295.
75. Emsley, P. a. C., K. (2004) Coot: model-building tools for molecular graphics, *Acta Cystallographica D60*, 2126-2132.

76. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V.B., Davis, I.W., et al. (2010) PHENIX : a comprehensive Python-based system for macromolecular structure solution, *Acta Cystallographica D66*, 213-221.
77. Adams, P. D. e. a. (2004) Recent developments in the PHENIX software for automated crystallographic structure determination, *Journal of Synchrotron Radiation 11*, 53-55.
78. Beillard, E. a. W., N.O. . (2005) Unraveling Kinase Signalling Pathways with Chemical Genetic and Chemical Proteomic Approaches, *Cell Cycle 4*, 434-437.
79. Wong, S. a. e. a. (2004) Sole BCR-ABL inhibitor is insufficient to eliminate all myeloproliferative disorder cell populations *PNAS 25*, 25.
80. Bishop, A. C., Buzko, O. and Shokat, K.M. (2001) Magic bullets for protein kinases, *Trends Cell Biol 11*, 167-172.
81. Laszlo, G., James, S.D.,Jacob P. T., Maria, B., and David, W. L. (2011) Unbiased Functional Proteomics Strategy for Protein Kinase Inhibitor Validation and Identification of bona fide Protein Kinase Substrates: Application to Identification of EEF1D as a Substrate for CK2, *Journal of Proteome Research 10*, 4887–4901.
82. Copeland, R. A. (2000) Enzymes, *Wiley 2nd Edition*, 137.
83. Thomas, J. A., and Koshland Jr, D.E. . (1960) Competitive Inihibition by Substrate during Enzyme Action. Evidence for the Induced-fit Theory, *Journal of American Chemical Society 82*, 3329-3333.
84. Johnson, M., Correia, J.J., Yphantis,D.A., and Halvorson, H.R. (1981) ANALYSIS OF DATA FROM THE ANALYTICAL ULTRACENTRIFUGE BY NONLINEAR LEAST SQUARES TECHNIQUES, *Biophysical Journal 36*, 575-588.
85. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state *Journal of Molecular Biology 372*, 774-797.
86. Krissinel, E. (2010) Crystal contacts as nature's docking solutions, *Journal of Computational Chemistry 31*, 133-143.
87. Angermayr, M., Hochleitner,E., and Lottspeich,F., and Bandlow, W. (2007) Protein kinase CK2 activates the atypical Rio1p kinase and promotes its cell-cycle phase-dependent degradation in yeast, *FEBS Journal 274*, 4654-4667.
88. Derewenda, Z. S. (2004) Rational Protein Crystallization by Mutational Surface Engineering, *Structure 12*, 529-535.
89. Dale, G. E., Oefner, C., and D'Arcy, A. (2003) The protein as a variable in protein crystallization, *Journal of Structural Biology 142*, 88-97.
90. Cooper D.R. , B., T.,Grelewska,K., Pinkowska,M.,Sikorska,M., Zawadzki,M., and Derewenda,Z. (2007) Protein crystallization by surface entropy reduction: optimization of the SER strategy, *Acta Crystallographica D63*, 636-645.
91. Longenecker, K. L., Garrard, S.M., Sheffield,P.J., and Derewenda,Z.S. (2001) Protein crystallization by rational mutagenesis of surface residues: Lys to Ala mutations promote crystallization of RhoGDI, *Acta Crystallographica D57*, 679-688.

92. Goldschmidt, L., Cooper,D., Derewenda,Z., and Eisenberg,D. (2007) Toward rational protein crystallization: A Web server for the design of crystallizable protein variants, *Protein Science 16*, 1569-1576.
93. Terwilliger, T. C., Adams,P. D. , Read, R. J. , McCoy,A. J., Moriarty, N. W. , Grosse-Kunstleve, R. W.,Afonine,P. V. , Zwart,P. H.,and Hung,L.-W. (2009) Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. , *Acta Crystrallographica*, 582-601.
94. Grosse-Kunstleve, R. W., and Adams, P.D. . (2003) Substructure search procedures for macromolecular structures, *Acta Crystallographica D59*, 1966-1973.
95. McCoy, A. J., Grosse-Kunstleve,R.W., and Adams, P.D. (2007) Phaser crystallographic software, *Journal of Applied Crystallography 40*, 658-674.
96. Terwilliger, T. C. (2000) Maximum likelihood density modification, *Acta Crystallographica D56*, 965-972.
97. Terwilliger, T. C., Grosse-Kunstleve, R. W. ,Afonine, P. V. ,Moriarty, N. W. , Zwart,P. H., . Hung,L.W., Read, R. J. ,and Adams,P. D. (2008) Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard, *Acta Crystallographica D64*, 61-69.
98. Terwilliger, T. C. (2003) Automated main-chain model building by template matching and iterative fragment extension, *Acta Crystallographica D59*, 38-44.
99. Terwilliger, T. C. (2003) Automated side-chain model building and sequence assignment by template matching, *Acta Crystallographica D59*, 45-49.
100. Terwilliger, T. C. (2003) Improving macromolecular atomic models at moderate resolution by automated iterative model building, statistical density modification and refinement, *Acta Crystallographica D59*, 1174-1182.
101. Dolin, P. J., Raviglione, M.C., and Kochi, A. (1994) Global tuberculosis incidence and mortality during 1990-2000, *Bulletin of World Health Organization 72*, 213-220.
102. (2010) The Global Tuberculosis Epidemic, *U.S. Global Health Policy Fact Sheet*.
103. (2010) Multidrug and extensively drug-resistant TB (M/XDR-TB): 2010 Global Report on Surveillance and Response., *WHO*.
104. Charity, J. C., Katz, E., and Moss, B. (2006) Amino acid substitutions at multiple sites within the vaccinia virus D13 scaffold protein confer resistance to rifampicin., *Virology 359*, 227-232.
105. Sturgill-Koszycki, S., Schaible, U. E., and Russell, D.G.(1996) Mycobacterium-containing phagosomes are accessible to early endosomes and reflect a transitional state in normal phagosome biogenesis., *EMBO J. 15*, 6960-6968.
106. Via, L. E., et al. (1998) Effects of cytokines on mycobacterial phagosome maturation, *Journal of Cell Science 111*, 897-905.
107. Russell, D. G. (1998) Cytokine activation leads to acidification and increases maturation of Mycobacterium aviumcontaining phagosomes in murine macrophages, *Journal of Immunology 160*, 1290-1296.

108. Russell, D. G. (2007) Who puts the tubercle in tuberculosis?, *Nature Reviews Microbiology 5*, 39-47.
109. Giacomini, E., Iona,E., Ferroni, L., Miettinen, M., Fattorini, L., Orefici, G., Julkunen, I., and Coccia, E.M. (2001) Infection of Human Macrophages and Dendritic Cells with Mycobacterium tuberculosis Induces a Differential Cytokine Gene Expression That Modulates T Cell Response1, *Journal of Immunology 166*, 7033-7041.
110. MacGurn, J. A., and Cox, J.S. (2007) A Genetic Screen for Mycobacterium tuberculosis Mutants Defective for Phagosome Maturation Arrest Identifies Components of the ESX-1 Secretion System, *Infection and Immnunity 75*, 2668-2678.
111. Gao, L.-Y., Pak, M., Kish,R., Kajihara, K., and Brown, E.J. . (2006) A Mycobacterial Operon Essential for Virulence In Vivo and Invasion and Intracellular Persistence in Macrophages, *Infection and Immnunity 75*, 1757-1767.
112. Stanley, S. A., Raghavan, S., Hwang, W.W., and Cox, J.S. (2003) Acute infection and macrophage subversion by Mycobacterium tuberculosis require a specialized secretion system, *PNAS 100*, 13001-13006.
113. Xu, J., Laine, O., Masciocchi, M., Manoranjan, J., Smith,J., Du, S.J., Edwards, N., Zhu, X., Fenselau, C., and Gao, L-Y. (2007) A unique Mycobacterium ESX-1 protein co-secretes with CFP-10/ESAT-6 and is necessary for inhibiting phagosome maturation, *Molecular Microbiology 66*, 787-800.
114. McLaughlin, B., Chon, J.S., MacGurn, J.A., Carlsson, F., Cheng, T.L., Cox, J.S., and Brown, E.J. . (2007) A Mycobacterium ESX-1–Secreted Virulence Factor with Unique Requirements for Export, *PloS Pathogen 3*, 1051-1061.
115. Gao, L.-Y., Guo, S., McLaughlin, B., Morisaki, H., Engel, J., and Brown, E.J. (2004) A mycobacterial virulence gene cluster extending RD1 is required for cytolysis, bacterial spreading and ESAT-6 secretion, *Molecular Microbiology 53*, 1677-1693.
116. Guinn, K. M., Hickey, M.J., Mathur, S.K., Zakel, K.L., Grotzke, J.E., Lewinsohn, D.M., et al. (2004) Individual RD1-region genes are required for export of ESAT-6/CFP-10 and for virulence of Mycobacterium tuberculosis, *Molecular Microbiology 51*, 359-370.
117. Pym, A. S., Brodin, P., Brosch, R., Huerre, M., and Cole, S.T. (2002) Loss of RD1 contributed to the attenuation of the live tuberculosis vaccines Mycobacterium bovis BCG and Mycobacterium microti., *Molecular Microbiology 46*, 709-717.
118. Behr, M. A., Wilson, M.A., Gill, W.P., Salamon, H., Schoolnik, G.K., Rane, S., and Small, P.M. (1999) Comparative genomics of BCG vaccines by whole-genome DNA microarray, *Science 284*, 1520-1523.
119. Ohol, Y., Goetz, D.H., Chan,K., Shiloh, M.U., Craik,C.S., and Cox, J.S. (2010) Mycobacterium tuberculosis MycP1 Protease Plays a Dual Role in Regulation of ESX-1 Secretion and Virulence, *Cell Host and Microbe 7*, 210-222.

120. Tan, T., Lee, W.L., Alexander, D.C., Grinstein, S., and Liu, J. (2006) The ESAT-6/CFP-10 secretion system of Mycobacterium marinum modulates phagosome maturation, *Cellular Microbiology 8*, 1417-1429.
121. Derrick, S. C., and Morris, S.L. (2007) The ESAT6 protein of Mycobacterium tuberculosis induces apoptosis of macrophages by activating caspase expression, *Cellular Microbiology 9*, 1547-1555.
122. Hsu, T., Hingley-Wilson, S.M., Chen, B., Chen, M., Dai, A.Z., Morin, P.M., Marks, C.B., Padiyar, J., Goulding, C., Gingery, M., et al. (2003) The primary mechanism of attenuation of bacillus Calmette-Guerin is a loss of secreted lytic function required for invasion of lung interstitial tissue, *PNAS 100*, 12420-12425.
123. Van der Wel, N., Hava, D., Houben, D., Fluitsma, D., van Zon, M., Pierson, J., et al. (2007) M. tuberculosis and M. leprae translocate from the phagolysosome to the cytosol in myeloid cells, *Cell 129*, 1287-1298.
124. Fortune, S. M., Jaeger,A., Sarracino, D.A., Chase, M.R., Sassetti, C.M., Sherman, D.R., Bloom, B.R., and Rubin, E.J. . (2005) Mutually dependent secretion of proteins required for mycobacterial virulence, *PNAS 102*, 10676-10681.